DESIGN, SCALING AND RELIABILITY OF DEVICES FOR HIGH-PERFORMANCE MIXED-SIGNAL APPLICATIONS

A Dissertation
Presented to
The Academic Faculty

By

Partha Sarathi Chakraborty

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
May 2015

DESIGN, SCALING AND RELIABILITY OF DEVICES FOR HIGH-PERFORMANCE MIXED-SIGNAL APPLICATIONS

Approved by:

Dr. John D. Cressler, Advisor
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Eric M. Vogel
Professor, School of Material Science and Engineering
Georgia Institute of Technology

Dr. Jeffrey A. Davis
Associate Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Azad J. Naeemi
Associate Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Saibal Mukhopadhyay
Associate Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: February 11, 2015

This work is dedicated to

my Parents, for their unconditional love and sacrifice;

my Wife, for her continuous love and support;

and

my Mentors, for generously sharing their valuable wisdom.

ACKNOWLEDGMENTS

I want to sincerely thank my advisor, Dr. John D. Cressler, for his invaluable mentoring, support, kindness, and encouragement. It was indeed an opportunity of a lifetime to learn from him.

My heartiest gratitude to Dr. Jeffrey A. Davis and Dr. Azad J. Naeemi for their time, flexibility, discussions, and for agreeing to be on my dissertation reading committee. My sincere thanks to Dr. Eric M. Vogel and Dr. Saibal Mukhopadhyay for agreeing to serve in my dissertation committee, and for sharing their valuable time and advise.

I gratefully acknowledge Semiconductor Research Corporation, Texas Instruments, IBM Microelectronics, Jazz Semiconductor, and IHP Microelectronics for their funding and support of this research work.

I owe it to the SiGe Devices and Circuit Research team for all their help and support; including my current and former team members for their collaboration and friendship. Special thanks to the SiGe Lab managers and station leaders who work to make the lab an awesome place for exploring.

A very special note of gratitude to Adilson Cardoso for all his time and contributions. I am particularly thankful to Anup Omprakash, Brian Wier, Uppili Raghunathan, Zelei Guo, Nedeljko Karaulac, Tiku Gebremariam, Yike Hu, Nelson Lourenco, and Michael Oakley for their friendship, support, discussions, and collaboration. My sincere appreciation goes to Anup, Adilson, and Michael Kroger for proof-reading sections of my dissertation.

I must mention some of my former lab members who had helped me when I started my research. Special thanks goes to Akil, Laleh, Marco, Prabir, Kurt, Nand, Stephen, Aravind, Ryan, Tom, Curtis, and Jiahui for their time and interactions.

Finally, I want to thank those special people in my life without whose strong support and love, it would not have been possible to even think about venturing into this work. To my wife Tanusree; for her patience, love, and continual support, else I could not have

embarked on this long journey. To my parents who brought me into this world, gave me the best education and life possible, and for their unconditional support, sacrifice, and encouragement at every important step even when we are thousands of miles apart. To my brother for his help. To my in-laws for their valuable and timely support. To all my mentors, friends, and well-wishers, particularly those from Georgia Tech, Arizona State, IIT Kanpur, Atlanta, Dallas, and Tempe who have been constantly there for us in this journey with all their heart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The development and continuous scaling of silicon-based process technologies has enabled highly integrated mixed-signal systems and circuits that incorporate analog, high-frequency, and digital circuit components to build cost-effective system-on-a-chip (SoC) solutions for high-performance mixed-signal applications. While emerging applications provide great impetus for aggressive scaling of the transistor performance, sustaining such an effort would require extremely meticulous attention to several important factors concurrently, such as yield, device mismatch, reliability, etc. Thus, predictive modeling of performance and reliability for developing these mixed-signal device technologies is required to sustain continuous performance scaling at a device, circuit, and system level in a cost-effective manner. Silicon-Germanium (SiGe) bipolar and complementary metal-oxide-semiconductor (BiCMOS) technologies in particular are positioned as an excellent candidate to satisfy all of these requirements. State-of-the-art SiGe BiCMOS technologies leverage deep-submicron silicon CMOS devices with bandgap-engineered SiGe heterojunction bipolar transistors (HBTs) in a single process technology that is suitable for a wide variety of high-performance, highly-integrated mixed-signal applications (e.g., system-on-chip (SOC), system-in-package (SiP)).

The objective of this research is to investigate and gain new understanding on how SiGe HBT device design couples with both performance scaling and reliability for mixed-signal applications (high-frequency and analog), and using this knowledge to enhance predictive modeling of performance and reliability for these devices. The effort is to develop a predictive device modeling methodology and simulation framework that can be used to design new mixed-signal device technologies and assess their performance and reliability concurrently. Ultimately, the goal is to highlight the need for device performance and reliability in a circuit environment, and establish best practices for practical modeling of these constraints. To support this objective, several specific areas were targeted to fill the existing

gaps in knowledge. This includes developing a technology computer-aided-design (TCAD) based integrated framework and methodology to study performance scaling and reliability in complementary SiGe HBTs; identifying factors determining the predictive nature of the simulated device figures-of-merit (FoM), studying the electrothermal constraints for scaling SiGe HBTs on thick-film silicon-on-insulator (SOI) to understand its impact on the DC and RF safe-operating-area (SOA) for the device, and performing reliability studies of hot-carrier damage and annealing in *npn* and *pnp* SiGe HBT devices in an effort to gain insight into the physical mechanisms involved and to develop fundamental understanding to aid TCAD modeling of hot-carrier damage in these devices. These studies provide the foundation for the bulk of this research, which addresses device-level performance and reliability modeling challenges through the application of a combination of existing and novel measurement methods; by using new device optimization strategies and modeling framework within TCAD. All of these individual studies resulting from the main research tasks are harmoniously tied together by a central theme: to develop a fundamental understanding about how the device design factors influence both performance scaling and reliability. Finally, this work provides methods for using the developed knowledge to enable predictive modeling of performance and reliability of SiGe HBT devices in BiCMOS technologies. Some of the key existing challenges and knowledge gaps are addressed by analyzing and reconciling the experimental data with simulation results.

The major contributions of this dissertation are summarized as follows:

- An improved optimization strategy and integrated TCAD framework for predictive modeling of the performance scaling complementary SiGe HBTs is presented. This includes the first feasibility study of a 200 GHz SiGe HBT and the first demonstration of a complementary SiGe HBT Roadmap using TCAD. A method of optimizing complementary SiGe HBT devices for performance matching over a range of bias conditions is also demonstrated [1, 2].

- The electrothermal constraints arising out of scaling vertical SiGe HBTs on thick-film SOI, and how these constraints impact the *dc* and RF SOA of the device is studied for the first time. New proposed measurement techniques are used with existing characterization techniques along with TCAD simulations to study several *dc* and RF FoM for the devices, and thus providing insight into the tradeoffs for scaling on thick-film SOI [3] [results to be submitted for publication].

- Factors influencing the predictive nature of the simulated output conductance of SiGe HBTs, and how those change with device performance scaling are identified, and assertions are verified through experimental results [4].

- The operative tunneling mechanisms that determine the reverse-biased emitter-base junction tunneling current in SiGe HBTs, and how this mechanisms are dependent on performance scaling are studied for the first time. Measurements and device simulations are compared for different generations of *npn* and complementary SiGe device technologies to provide insight into the predictive modeling of the tunneling current. Reliability degradation due to hot-carriers generated by the reverse-biased tunneling stress on the emitter-base junction and its dependence on performance scaling is also studied for the first time [5].

- The first experimental proof for the physical mechanisms involved in the hot-carrier induced mixed-mode stress damage of *pnp* SiGe HBTs is presented. The differences in the type of hot-carrier driving the mixed-mode stress damage between *npn* and *pnp* devices are highlighted as part of this study [6].

- Demostration of the best cryogenic performance reported till date of a ”best-of-breed” *npn* SiGe HBT, and its comparison with the cryogenic performance of a *pnp* from the ”best-of-breed” complementary SiGe HBT technology are presented with insight into their implications. This work also presents the first study on cryogenic

performance of complementary SiGe HBTs [7] [results to be submitted for publication].

- New safe-operating area construction methods are proposed to compare impact-ionization coefficient and hot-carrier reliability of devices more generally across different device technologies, stress methods, degradation criteria, and device geometries in an integrated way. These results have strong implications for use in the device technology development cycle. In addition, the first experimental validation of the reaction-diffusion (R-D) model for hot-carrier damage operative in SiGe HBTs are presented using evidence from case studies for the annealing of hot-carrier damage [results to be submitted for publication].

CHAPTER 1

INTRODUCTION

The invention of the semiconductor transistor and integrated-circuit triggered an unmatched revolution in commercial electronics, telecommunications, and computers, which continues to date. As silicon based processing technologies for fabricating transistors and integrated-circuits matured in the next phase, it triggered incessant device and system performance scaling.

Initially, the stimulus for performance scaling was provided by their potential implementation into existing applications like radio communications, and development of new applications like computer mainframes. However, as silicon based solid-state products scaled and proliferated into every imaginable aspect of human life, the transistor became an ubiquitous part of our daily life.

The silicon based semiconductor products revolutionized the world like never in the history of human civilization, specifically with the advent of scaled transistors for faster digital applications enabling very high performance computational resources. Scaling of the silicon based semiconductor transistor for digital applications has continued ever since. More recently, the learnings from digital silicon technologies have been leveraged for the scaling of non-digital semiconductor technologies, like analog, mixed-signal, high-frequency (RF and mm-wave), etc. This again has led to the proliferation and use of a wide range of solid-state devices including bipolar transistors, CMOS transistors, MEMS devices, diodes, integrated passives, etc. for realizing integrated systems with a wide range of functionality, potentially on the same die or package with an ever decreasing form factor.

Development of silicon-germanium BiCMOS process has in many ways enabled the possibility of highly-integrated system-on-chip (SoC) geared towards high-performance mixed-signal applications. In these target applications, the circuit-level metrics are closely

tied to device level performance and reliability metrics. Hence, a fundamental understanding of the factors influencing device level performance and reliability is essential towards improving the circuit and system level metrics. An in-depth study and analysis of device performance and reliability would demand a strong knowledge of the fundamental physical processes operative within the device. However, the broad range of potential mixed-signal applications and design considerations leads to the devices and circuits encountering unforeseen performance and reliability issues. These issues makes studying the SiGe BiCMOS process technologies and the SiGe HBT devices in particular a continuous and worthwhile research pursuit.

1.1 Research Objectives

The objective of the research presented in this dissertation is to investigate and gain new understanding on how device design simultaneously couples with both performance scaling and reliability for mixed-signal applications (RF and analog), and applying this developed knowledge to enhance predictive modeling of performance and reliability for these devices. The research targeted several specific areas to support this objective, and are listed as follows:

1. Provide an improved optimization strategy and integrated TCAD framework for predictive-modeling of the performance scaling in complementary SiGe HBTs.

2. Identify factors affecting the predictive nature of the simulated FoM for SiGe HBTs like the output conductance.

3. Use predictive modeling to investigate the impact of the reverse-biased emitter-base junction tunneling currents and how that couples to performance scaling.

4. Provide insight into the operating-constraints arising out of the electrothermal effects resulting from scaling conventional SiGe HBTs on thick-film SOI (including effects on ac, dc characteristics and RF linearity).

5. Present a proof for the physical mechanisms involved in the mixed-mode stress damage of *pnp* SiGe HBTs.

6. Investigate the cryogenic performance of best-of-breed *pnp* SiGe HBTs in a complementary SiGe BiCMOS technology, and compare its cryogenic performance to the best-of-breed *npn* device.

7. Develop methods for comparison of the *npn* and *pnp* SOA in a complementary SiGe BiCMOS technology, and for understanding the annealing of hot-carrier damage.

The results from these studies are tied together with a overarching theme, i.e., to develop a holistic understanding about how the device design factors influence the performance scaling and reliability of these devices, and using this newly established knowledge towards predictive modeling of performance scaling and reliability in mixed-signal (analog and RF) devices for BiCMOS technologies. The presented research topics remain in the device realm and fill some of the existing gaps in the understanding of predictive-modeling for performance scaling and reliability. This knowledge can be leveraged by device designers to predictively model both performance and reliability in an integrated TCAD framework for new technology development.

1.2 Thesis Outline

The structure and organization of the thesis is as follows:

- A background of the growth of BiCMOS technologies for mixed-signal applications, motivation for studying complementary SiGe technology, and an in depth literature review of a priori knowledge are presented in Chapter 2.

- In Chapter 3, a new fully integrated predictive simulation framework for complementary SiGe HBT device design is introduced for the first time, and is useddd with a device design methodology to investigate the scaling of these devices. This TCAD

framework and methodology allows for the first time an investigation into the feasibility of scaling C-SiGe HBTs up to the 200 GHz node, and is used for TCAD studies the following chapters as well [1, 2].

- Characterization of the operating constraints resulting from electro-thermal effects of scaling SiGe HBTs on thick-film SOI, and how it influences the SoA of the device is presented in Chapter 4, [3].

- A study on the predictive output conductance modeling of SiGe HBTs in TCAD based on the aforementioned simulation framework in conjunction with device measurements is presented in Chapter 5 [4].

- In Chapter 6, the developed TCAD modeling framework is used to study how SiGe HBT performance scaling influences the reverse-biased EB junction tunneling mechanisms. Reliability of the devices from hot-carriers generated under stress in the tunneling dominated regime is studied, and how it couples with scaling is highlighted at the end. [5].

- In Chapter 7, experimental results from mixed-mode stress on *pnp* devices in a C-SiGe BiCMOS technology are analyzed to provide the first experimental proof and insight into the underlying physical mechanisms involved in the hot-carrier damage processes, while highlighting differences with the *npn* devices [6].

- In Chapter 8, cryogenic performance of a best-of-breed *npn* SiGe HBT is presented and compared it to that of a best-of-breed *pnp* SiGe HBT in a complementary SiGe BiCMOS technology. As part of this study, a record cryogenic performance of *npn* SiGe HBTs is presented [7]

- In Chapter 9, first an investigation on the methods to construct and compare the SOA for comparing reliability of *npn* and *pnp* SiGe HBTs in a complementary SiGe BiCMOS technology is presented. Some new methods of SOA contruction for comparing

device reliability are proposed. In the second part, a detailed study on the annealing mechanisms of the hot-carrier induced damage is presented.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Growth of BiCMOS Technologies

The phenomenal evolution and growth of the wireless and wireline communications industry in recent years has re-ignited the need for high performance BiCMOS (bipolar-complementary-metal-oxide-semiconductor) technology platforms. This trend has been driven by the fact that whereas CMOS still forms the backplane for the digital part of mixed-signal systems, bipolar devices are better suited for implementing the high-frequency and analog/mixed-signal (AMS) blocks. This opportunity gives the silicon-germanium (SiGe) heterojunction bipolar transistor (HBT) an edge over its rivals, because of its fabrication compatibility with conventional Si CMOS processing and its potential for extremely high levels of performance and integration. Consequently, SiGe HBTs have seen significant proliferation into mixed-signal applications leading to its performance scaling and evolution through multiple generations of BiCMOS technologies [8–10]. Therefore, in a way highly-integrated, performance-constrained mixed-signal systems, geared towards system-on-chip (SoC) and system-in-package (SiP) applications drives the need for Si-based BiCMOS technologies. This typically comes within the "More-than-Moore" domain as defined by the International Technology Roadmap for Semiconductors (ITRS) [11]. Evolution of the SiGe BiCMOS technologies can be classified into two broad categories. One driven by *npn*-only bipolar devices geared primarily towards high-frequency microwave and mm-wave applications; and the other driven by complementary bipolar devices (with both the *npn* & *pnp*) geared primarily for high-performance analog and mixed-signal applications.

2.2 SiGe BiCMOS Technology for Mixed-Signal Applications

The performance scaling of SiGe HBTs continues to strongly push the roadmap for advanced BiCMOS process technologies. SiGe BiCMOS platforms are making rapid inroads in performance-constrained mixed-signal integrated circuits across many application domains. Sustained scaling clearly requires the successful demonstration of SiGe HBTs with both exceptional performance and robust reliability, followed by their implementation into various successful commercial products [8, 9].

While device designs are constrained by the challenges in fabrication technologies, the pace of the performance roadmap development and implementation is driven by the market economics of the target products, as expected. For example, currently the development of high-speed *npn*-only BiCMOS technologies is primarily tied to emerging millimeter-wave application markets. It is difficult to have initial estimates of the product volumes related to these markets. Firstly, because most of these products deal with new applications, the development of which depends on their marketing success. Secondly, because at least some if not all of these markets may also be covered by competing CMOS and other semiconductor technologies. In the context of these new products and applications, potential markets with sufficient volumes alone can justify the development of new BiCMOS technologies. Development of a new BiCMOS technology in a way requires predictive estimation of device performance and reliability to begin with, followed by the validation of reliable operation for both the fabricated devices and their applications. In the future, other semiconductor and new emerging device technologies (e.g., graphene) are going to become increasingly important as competitors for SiGe BiCMOS technologies. As an example, graphene is being recently explored for its suitability as an alternate technology for high-frequency applications under the "More-than-Moore" domain [11].

The purpose of the research presented here encompasses several broad objectives. Firstly, to develop and use predictive device modeling tools to design a roadmap for complementary SiGe (C-SiGe) HBTs and use it to estimate their performance and reliability metrics.

New and existing experimental data will be used wherever possible to validate these tools and understand their implications. Secondly, it attempts to study the feasibility and implications of scaling SiGe HBTs on thick-film (TF) silicon-on-insulator (SOI) substrates. Thirdly, this research attempts to study the fundamental mechanisms behind hot-carrier induced degradation of C-SiGe HBTs under mixed-mode stress. Finally, cryogenic characterization of the highest performance *pnp* SiGe HBTs available commercially is performed and its implications for both *pnp* and C-SiGe HBT scaling are summarized.

2.3 Literature Review

This literature survey covers the spectrum of important history in each topic through the cutting-edge.

2.3.1 Scaling of Complementary SiGe HBTs for Mixed-Signal Applications

Complementary bipolar technology [*npn* + *pnp* bipolar junction transistors (BJTs)] has been long considered the "gold standard" for analog applications requiring flexibility in designing circuits with high speed, low power, low noise, high linearity, large bandwidth, large voltage swing, and output drive [12–15]. Introducing graded Ge into the base of the Si BJTs to build band-gap-engineered C-SiGe HBTs can provide significant leverage in the tradeoffs involved in designing *npn* and *pnp* BJTs with matched performance (e.g., comparable f_T) and reliability [16, 17]. However, the inherent minority-carrier transport issues that stem from the Ge-induced valence-band offset and lower minority carrier mobility associated with the *pnp* SiGe HBT device design need to be carefully addressed to successfully scale C-SiGe HBT technology for high-frequency operation [18–25].

An earlier technology-computer aided design (TCAD) based study of C-SiGe HBTs was limited to drift-diffusion transport models, one-dimensional (1-D) hypothetical doping profiles, and devices with much lower peak f_T and f_{MAX} [26]. This earlier study gave no consideration to the SiGe layer stability or the impact of simultaneously changing multiple device parameters. The accuracy of simulated device figures-of-merit (FoM) needs to be

compared with experimental devices to assess the predictive nature of the study. Although there have been studies on self-heating (SH) effects in trench-isolated SiGe HBTs on SOI (since SOI provides distinct advantages for C-SiGe scaling) [27, 28], no studies of vertical SiGe HBTs on TF SOI substrates with a performance of peak $f_T > 100$ GHz have been reported yet. The main focus of performance scaling has been with the use of bulk-silicon substrates.

There are traditional technological challenges facing the *npn*-only portion of the roadmap (that focuses on the development of new *npn* device architectures and their consequent integration with advanced CMOS nodes). These challenges are different from the ones encountered by the complementary (*npn* + *pnp*) portion of the scaling roadmap. Complementary bipolar technology development focuses on control of the vertical profile of the *pnp* transistors to achieve a performance comparable with the *npn* devices, and their subsequent co-integration with the *npn* transistors and advanced CMOS devices. These different *npn* and *pnp* device design challenges may impact the evolution of their respective performances such as the trade-offs between f_T, f_{MAX} and the breakdown voltages. However, these challenges are probably not the main show-stoppers today. While device design and fabrication challenges still exists, development and implementation of new manufacturable BiCMOS process technologies into viable products will be driven by their market for existing and emerging applications; coupled with economy of scale.

2.3.2 Predictive TCAD Modeling of SiGe HBTs

SiGe BiCMOS technologies are now playing an increasingly significant role in high performance analog and mixed-signal (AMS) applications [8, 9]. Several AMS circuits require high $f_T \times BV_{CEO}$ and $\beta \times V_A$ products for the SiGe HBT-based building blocks. Optimization of the relevant device FoM for any new process technology requires TCAD for successful predictive modeling. This procedure entails careful use of the correct transport and physical models, with calibrated input parameters. Although there have been extensive studies on the accurate TCAD-based modeling of "standard" SiGe HBT FoM such as f_T, f_{MAX},

and BV_{CEO}, very few studies have focused on the output conductance or Early voltage (i.e., V_A). Another important aspect that remains to be understood is how this TCAD modeling challenge couples to device performance scaling.

2.3.3 Scaling SiGe HBTs on Thick-Film SOI

The critical need for an SOI-based BiCMOS platform stems from the fact that the substrate needs to be isolated to decouple the noise generated in the digital part of the system from the on-chip radio-frequency (RF) and analog blocks. Additionally, SOI offers benefits derived from fully isolated devices, reduced parasitics, and reduced sensitivity to radiation-induced single-event upset (SEU). SOI-based BiCMOS technologies further enable the integration of complementary bipolar (*npn* + *pnp*) and CMOS on a single chip [8–10].

As SiGe HBTs continue the path of performance scaling through advanced lithography nodes and superior process innovations, the current density (J) and the peak electric field (E) at the device junctions during device operation continue to scale upward, leading to increased power dissipation from Joule heating ($J \cdot E$) per unit area. To make this problem more challenging, thermal resistance (R_{TH}) of the SiGe HBTs continue to increase due to the use of advanced electrical isolation techniques (e.g., shallow-trench isolation (STI), deep-trench (DT), SOI) for performance scaling. These electrical isolation techniques result in increased thermal isolation. Consequently, this leads to increasing mismatch between the heat dissipation and heat conduction mechanisms, thus resulting in stronger SH effects and increased junction temperatures (T_J) during operation of the highly-scaled SiGe HBTs [29, 30].

TF SOI permits the use of the existing conventional vertical SiGe HBT device structure with a selectively-implanted collector and a highly-doped sub-collector. To date there have been several demonstrations of SiGe HBTs fabricated on SOI [31–35]; however, there are no reported devices with both peak $f_T/f_{MAX} > 100$ GHz on TF SOI. While *npn* SiGe HBTs with both peak $f_T/f_{MAX} > 300$ GHz performance exists on bulk-silicon, the performance of SiGe HBTs on TF SOI needs to be scaled beyond 100 GHz and assessed for associated

SH effects. This will provide valuable understanding about the implications of scaling the performance of SiGe BiCMOS technologies on TF SOI. SH induced instabilities resulting from strong ET feedback in a device have been earlier experimentally studied and analyzed in both single and multi-finger III-V and SiGe HBTs [36–39]. However, SH effects of scaling conventional SiGe HBTs on TF SOI needs to be further investigated to fully understand the electrothermal constraints governing performance scaling of TF SOI based BiCMOS technologies beyond the 100 GHz barrier.

2.3.4 Emitter-Base Junction Tunneling Current in SiGe HBTs

Scaling of bipolar transistors (BJTs or HBTs) requires higher doping within certain regions of the device. High base doping, for instance, is required to prevent punch-through, lower base resistance, and increase output impedance; whereas higher emitter doping leads to higher emitter efficiency and lower emitter series resistance. The Ge profile in SiGe HBTs is leveraged to partially relax some of these heavy doping requirements. However, when the base doping exceeds a certain level, the zero-bias (ZB) emitter-base junction (EBJ) space-charge region (SCR) narrows considerably, resulting in a sufficiently small barrier for efficient carrier tunneling between the conduction and valence bands. The result is a non-negligible current from direct band-to-band tunneling (BTBT) and trap-assisted tunneling (TAT) mechanisms in low forward-bias (FB) through reverse-bias (RB) of the EBJ [40–43]. This excess tunneling current will pose a fundamental limitation on performance scaling since it sets a lower limit on the reverse leakage and an upper limit on the achievable EBJ breakdown voltage. These tunneling constraints clearly impact device reliability under reverse-biased stress. BJTs and HBTs used in BiCMOS mixed-signal integrated circuits are often subjected to EBJ reverse-biasing during normal circuit operation, which inevitably leads to a reliability concern for long-term operation [44].

As the performance of SiGe HBTs continues to scale, the base doping can easily reach $10^{19}\ cm^{-3}$ or more in state-of-the-art devices. At these high doping levels, the ZB peak electric field ($E_{0,\text{peak}}$) at the EBJ will exceed 10^6 V/cm; hence, the tunneling component of the

base current (I_B) under reverse-bias becomes increasingly important for such SiGe HBTs. Both $E_{0,peak}$ and tunneling current of the EBJ needs to be minimized during device optimization through predictive TCAD modeling of the scaled SiGe HBTs, as it significantly impacts the device reliability related to reverse-biased stress on the EBJ [43].

Historically, multiple physical models have been proposed to incorporate tunneling mechanisms in device TCAD simulators [40, 41, 45]. However, their effectiveness for predictive modeling of the tunneling current across multiple generations of SiGe HBT devices has not been investigated. The quantum-mechanical nature of the basic tunneling mechanism makes it computationally challenging to effectively include the models in device simulators with drift-diffusion (DD) or hydrodynamic (HD) transport models. Therefore, it is important for any predictive device simulation framework to model this mechanism accurately.

2.3.5 Mixed-Mode Reliability Mechanism in *pnp* SiGe HBTs

The availability of complementary bipolar transistors (*npn* + *pnp*) of comparable performance and reliability within a BiCMOS process technology offers significant advantages (e.g. high linearity, high speed, low power, large voltage swing, and output drive) when designing high-performance analog ICs compared to *npn*-only technologies [8, 9]. Despite a historical bias for the development of *npn* over *pnp* SiGe HBTs due to the inherent manufacturing complexity associated with fabricating complementary devices of comparable performance, two C-SiGe HBT BiCMOS process technologies were recently reported [18, 19]. Mixed-mode (MM) reliability has been recently used to assess the robustness of SiGe HBTs to high-voltage and medium-to-high current stress beyond the traditional safe-operating area (SOA) of the device [46]. While extensive investigations have been reported on the MM reliability of *npn* SiGe HBTs [47, 48], negligible data is available on the MM reliability of *pnp* SiGe HBTs [16, 17]. These recent investigations have examined MM reliability degradation mechanisms for *npn* SiGe HBTs [47, 48], providing

insight into the spatial location of the impact-ionization (I-I) induced damage inside the device structure. However, the explanation for the underlying physical mechanisms driving the damage process still remains open.

2.3.6 Cryogenic Study of State-of-the-Art *npn* and *pnp* SiGe HBTs

Recently reported cryogenic measurements on *npn* SiGe HBTs and TCAD simulations show that a viable scaling roadmap for *npn* SiGe HBT performance should be possible, albeit with associated non-trivial fabrication challenges. Due to its bandgap-engineered base, SiGe HBT performance is well-known to be enhanced by cooling, and thus temperature can be used as a "scaling-lever" to provide insight into the performance scaling for room temperature operation. There have been multiple cryogenic studies to date geared towards understanding the scaling of *npn* SiGe HBTs [49, 50], however similar studies have not been performed for state-of-the-art *pnp* SiGe HBTs. Comparison between the cryogenic performance scaling of state-of-the-art *pnp* devices with that of *npn* devices (either with different or comparable performance) and its implications are essential to understanding further scaling of C-SiGe devices.

2.3.7 Safe Operating Area Construction for Hot-Carrier Reliability and Annealing of Hot-Carrier Induced Damage

Scaling of SiGe HBT performance has led to device operation at increasingly shrinking voltage limits and higher current densities. Consequently, to leverage the maximum performance in a circuit environment, the devices are more often being pushed towards operation at or above the safe-operating area (SOA) boundaries which define the estimated lifetime of the device for reliable operation. The SOA limits are based on the time-dependent degradation of a relevant device figure-of-merit (FoM) derived from accelerated stress test methods. The accelerated stress methods are used to estimate reliable aging of the devices to safely meet the target lifetime for potential applications. Typically, the minimum BV_{CEO} (over process variations and temperature range of the application) is recommended to be the maximum usable voltage (or SOA) limit for reliable operation of all devices in that

technology, over the entire lifetime of the target applications. Apart from the voltage limit, other operating constraints are typically imposed by the minimum current density required for the onset of electromigration and the maximum allowed power dissipation (or self-heating). All of these limits are being continuously scaled down with aggressive device performance scaling. Defining the device SOA boundaries for any circuit application is a non-trivial problem, considering the various damage creation and annealing processes operative across the limits of the SOA on the output plane, with different underlying physical mechanisms and their dependence on temperature. The SOA limits are furthermore defined by the device geometry and target lifetime of the device in a circuit application. The device degradation and annealing mechanisms are typically driven by factors such as the local time-dependent electric field and current density within the device, lattice temperature, device geometry, device physical structure, and the process technology.

In essence, the concept of a single SOA cannot be generalized for all device geometries in any technology (considering the worst case scenario would be very conservative), as such a definition would ignore the fact that the electrical manifestation of the damage and the SOA boundaries are strongly dependent on the device geometry. Thus, this approach would leave untapped performance for device geometries with a more conservative SOA on the output plane. Thus a price of valuable lost performance has to be paid if conservative limits from device geometries with worst case damage estimates are used to generally define the SOA for all device geometries in that technology. Another concept that needs to be highlighted is the fact that the SOA boundaries or limits (including damage and annealing regions) are dependent on the accelerated stress-time and stress-method used to estimate the reliable lifetime of the device, in addition to the device geometry and electrical criterion used to measure degradation. In other words, the same degradation criterion or device geometry can generate a different SOA being measured from a different accelerated stress-method or stress-time. Conversely, the same stress-time and stress-method can generate a different SOA with the use of different degradation criterion or device geometry.

For example, SOA derived from the forward and inverse gummel leakage will typically show differences for the same device, as the physical location of the damage (oxide-silicon interfaces) within the device responsible for the performance degradation in both cases are distinctly different (emitter-base spacer vs. shallow-trench isolation). In addition to deriving a SOA from *dc* stress, a different rate of degradation and SOA will be observed when using RF stress in conjunction with *dc* stress as an accelerated stress mechanism to study RF reliability [51]. Furthermore, a SOA derived at room temperature can be quite different from that derived at either end of the target temperature range for the application, the differences being caused by the temperature dependence of the fundamental damage mechanisms determining the different SOA limits. These concepts are in stark contrast to how a fixed SOA has been earlier derived and reported for a single device technology from a single accelerated stress method, and then generalized for all stress-time and device geometries [52].

The SiGe HBT devices can be operated at different regions of the output plane as a function of time in various high performance RF and mixed-signal applications, thereby dynamically undergoing a complex combination of different damage and annealing mechanisms. Consequently, it can thus generate a sophisticated response dependent on the combination of different stress-conditions and stress-durations experienced across the output plane. Thus it becomes imperative to consider the bias dependence of damage creation and annealing processes to model and predict the degradation of a device being operated in a mixed-signal circuit environment. In one approach, a combination of empirical and physics based modeling was implemented for use in a compact model environment [53]. Recently, a fully physics-based approach using a TCAD environment has been proposed [54, 55], enabling time dependent estimation of damage creation and in-line incorporation back into the TCAD device model for a dynamic assessment of the device reliability with evolution of stress-time and conditions. This approach is very suitable for use in the predictive modeling of both SiGe device and circuit reliability, and is compatible for use with different

generations of device technology. This TCAD based approach uses a reaction-diffusion (R-D) model to predictively estimate the hot-carrier induced damage creation (or interface trap generation) process at the critical oxide-silicon interfaces within the device, including physics-based models for hot-carrier generation and propagation to the interfaces.

However, this TCAD-based method to estimate the hot-carrier degradation presently has some limitations in its implementation. Firstly, it only uses models for damage creation, in the absence of existing valid models for damage annealing. Secondly, it leads to a monotonic increase in damage with increasing stress-time in the absence of accurate boundary conditions, and does not in any way show saturation of the maximum damage created (say trap density, etc.). Both of these aspects can be implemented through careful consideration of the underlyng physical mechanisms involved and then using mathematical implementation of the model in the TCAD framework and post-processor.

CHAPTER 3

PREDICTIVE SCALING AND OPTIMIZATION OF COMPLEMENTARY SIGE HBTS

This study leverages technology computer-aided design (TCAD) and presents for the first time a comprehensive study of the device design challenges and optimization issues that are unavoidable in designing and scaling C-SiGe HBTs for high-performance analog and RF applications [2].

In this C-SiGe scaling study we utilize the commercially-available Sentaurus Workbench (SWB) TCAD environment and its entire suite of simulation tools [56] to investigate two-dimensional (2-D) SiGe HBT device structures on bulk Si, using both shallow-trench (STI) and deep-trench (DT) for isolation (Fig. 1). This approach provides a more relevant platform from a technology development perspective for the high-performance, mixed-signal semiconductor industry.

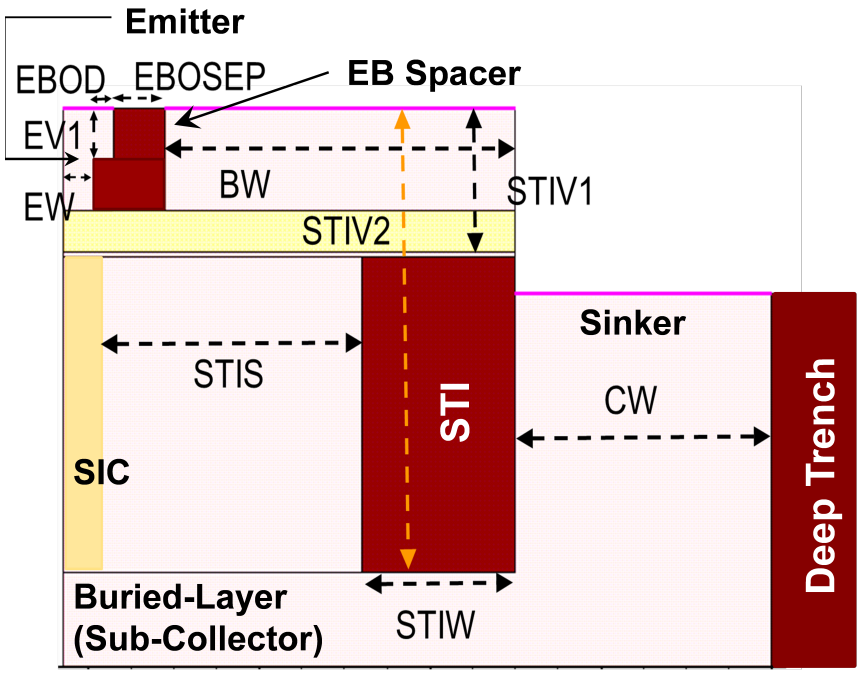

Figure 1: 2-D cross-section of the simulated and parameterized device structure. Parameters shown here were used for tuning the physical device structure (both vertical and lateral) [2].

An earlier initial study illustrated some of the issues associated with scaling C-SiGe HBTs towards the 200 GHz performance node [1]. In the present section, the previous work is expanded and employed to develop a C-SiGe HBT scaling roadmap. This work presents for the first time: a calibrated performance scaling study of C-SiGe HBTs; the development and benchmarking of a C-SiGe HBT scaling roadmap; addresses TCAD-based C-SiGe HBT predictive device modeling issues that one necessarily encounters in such a study; and introduces an optimization methodology that must be considered while simultaneously optimizing C-SiGe HBTs for use in both high performance and low power circuits.

3.1 TCAD Simulation Framework and Methodology

The doping and Ge profiles for both *npn* and *pnp* SiGe HBTs, along with both their vertical and lateral physical cross-sections, were carefully parameterized (Fig. 1). Full hydrodynamic carrier transport code, with Philips unified mobility models, and including Okuto-Crowell model for non-local impact ionization, bandgap narrowing, carrier recombination (Shockley-Read-Hall and Auger), and Canali model for high-field velocity saturation (using carrier temperature as the driving force), were used for the device simulations presented [1, 57–59].

The model parameter files were calibrated to measured data from a commercial 200 GHz *npn* SiGe HBT platform. Similar parameter files were used for both the *npn* and the *pnp* devices. The SiGe parameter file (available within SWB) used for this study interpolates between the Si and Ge properties based on the Ge mole fraction. Further accuracy in the modeling of strain effects on the carrier transport within the SiGe layer can be implemented using parameter files generated from separate Monte-Carlo (MC) simulations (for the *npn* and *pnp* devices) based on their individual doping and Ge profiles. However, this would still not account for effects of the Si cap layer, processing steps, and carbon (C) incorporation to suppress boron out-diffusion on the final SiGe layer strain. Important

parameters for the recombination models were carefully calibrated, since accuracy of these parameters is key to achieving predictive simulation of I_B and the *dc* current gain (β_{DC}), and hence BV_{CEO}.

The selectively-implanted collector (SIC) regions of the simulated devices were constructed using two Gaussian profiles, each of which can be tailored independently to emulate the individual SIC implants in a fabricated device. This allows for a greater flexibility in optimizing for the trade-off between the onset of Kirk effect (and heterojunction barrier effects) and the collector-base junction breakdown voltage (BV_{CBO}). The SIC region was designed to have a lateral straggle to emulate the doping profile that extends beyond the emitter window to the STI, as in a real device. Consequently, each of the individual SIC Gaussian profiles was parameterized for independent control of peak doping position and concentration, as well as vertical and lateral straggle [1]. There has been several earlier studies on TCAD- based device scaling and collector profile optimization of *npn* SiGe HBTs [60–62]; and earlier methods, wherever applicable, have been used in this investigation for designing the complementary devices.

A stability calculation of the SiGe layer (for any given doping and Ge profile) using the Matthews-Blakeslee criterion with Fisher's cap layer correction was integrated directly into the simulation environment [63]. This provides an estimation of the SiGe layer stability for each profile variation. All profiles considered in this study were thermodynamically stable. accounting for C incorporation would only improve the SiGe layer stability.

A pinched-base sheet resistance (R_{bi}) calculation was performed for each device profile. Strain effects were accounted for in this calculation based on the doping and Ge profiles, and the carrier mobilities [9]. Device simulations were performed to extract the standard *dc*, *ac*, and output characteristics (forced-V_{BE} and forced-I_B) for each profile. From the device simulation results, important figures-of-merit (FoM) (e.g., f_T, f_{MAX}, β_{DC}, BV_{CEO}, and V_A) were extracted. An integrated 1-D and 2-D quasi-static transit-time (QSTT) analysis was used to compare regional transit-times and fine tune the candidate device profiles based

on the limiting factors for *ac* performance and the onset of heterojunction barrier effects (HBE) in the device [60, 64–66].

Half of the device cross-section was simulated to utilize symmetry around the center of the emitter and thereby reduce the simulation time. Fig. 2 shows a flow diagram of the steps adopted in the device design and optimization methodology used for developing the profiles of C-SiGe HBTs in this study. All simulations were isothermal and performed at room temperature (300 K) unless otherwise noted. The target performance metrics were defined based on existing *npn*-only SiGe HBT device technologies. The *ac* performance of the optimized devices was simulated for different biasing modes (constant V_{CB} and V_{CE}) to ensure that they are comparable in performance over a broad range of operating conditions and design topologies (common-base or common-emitter). For example, a wide range of relevant V_{CB} and V_{CE} values were considered to simultaneously account for circuit designs ranging from low-power to high-performance applications. However, the same methodology can be used for TCAD-based matching studies of *npn* vs. *pnp* performance over temperature, provided valid parameter files are available over the entire temperature range. This was beyond the scope of the study.

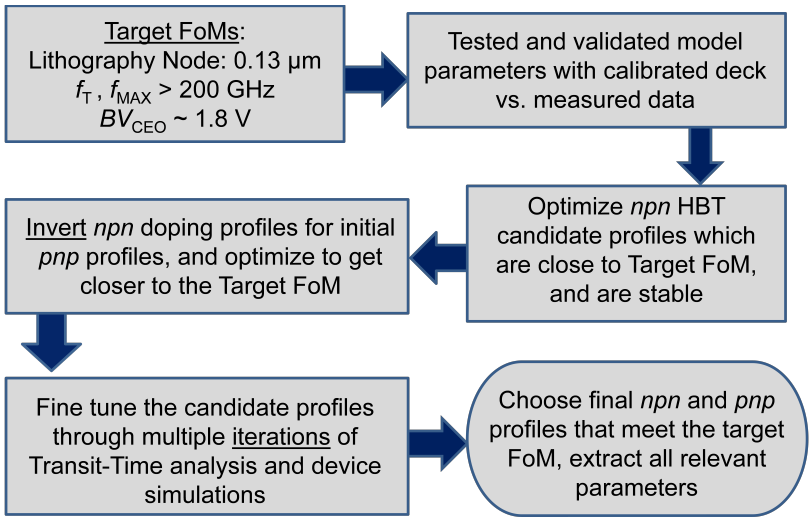

Figure 2: TCAD simulation steps used for the complementary SiGe HBT device scaling and optimization methodology (for any target performance node) [2].

3.2 Device Modeling Issues

Although most of the important device FoM such as f_T, f_{MAX}, β_{DC}, and BV_{CEO} were modeled reasonably well using our TCAD simulation framework under isothermal conditions, realistic V_A (Early voltage) values for the scaled devices could only be simulated by appropriately considering self-heating together with impact-ionization in the hydrodynamic device simulations [1, 4]. This resulted in longer simulation times and convergence issues. Greater accuracy in predictive V_A estimation can be achieved by using 3-D device structures with more accurate thermal boundary conditions in the device simulator. The divergence between the V_A simulated with and without self-heating clearly increases with J_C, device thermal resistance (R_{TH}) and the device performance, owing to stronger self-heating effects [3].

Lateral scaling of a candidate 200 GHz *pnp* device has shown that the peak f_T and maximum β_{DC} remains fairly stable, being solely dependent on the vertical profiles, while f_{MAX} scales with the evolving lithography node as R_{bi}, R_{bx}, C_{cb} and C_{cs} are dependent on the lateral device structure. Lateral scaling in the simulation decks was incorporated through changing both the intrinsic (emitter width or EW) and extrinsic (through emitter-base spacer width or EBOSEP) part of the base region, based on the lithography node considered (refer to Fig. 1). The 2-D lateral structure parameters were chosen such that the *npn* performance was reasonable when compared to reported values in the literature, but was kept fixed for the *npn* and the *pnp* devices at the same performance node. However, in general it needs to be understood that these parameters will play a role in determining f_{MAX} and self-heating of the device, and can be appropriately leveraged as additional tuning knobs for achieving comparable electro-thermal performance of the *npn* and *pnp* devices. The present study was mainly focused on vertical profile optimization of the devices, with the intent to shed light on the intrinsic profile differences of the *npn* vs. *pnp* at comparable performance and 2-D physical dimensions. Some of the important structural parameters of the optimized candidate devices are provided in Section 3.4.

Analysis of the excess carrier concentration vertically along the middle of the emitter shows that no well-defined quasi-neutral region exists within the base of a highly scaled 200 GHz *pnp* HBT (the same holds true for *npn* devices). This is critical in defining the quasi-neutral or space-charge regions within the metallurgical base for accurate estimation of the regional transit-times from the QSTT analysis [66].

HBE is a stronger constraint for optimizing *pnp* SiGe HBTs due to the inherently larger band offset for minority carrier transport across the device [26]. 1-D transit-time (T-T) components along the center of the emitter for a 200 GHz *pnp* profile shows that the emitter-base (EB) junction T-T (or the hole inverse velocity) component limits the device performance at low-to-moderate injection, whereas the onset of HBE at high-injection results in the collector-base junction T-T limits the performance of the device [1].

For the purposes of this study, the doping of the emitter, base and collector contact regions were kept comparable for the *npn* and *pnp* devices at a specific performance node. At comparable doping, the n-type doped regions will have lower resistance than the corresponding p-type doped regions due to higher majority carrier mobility. Thus, in this study emitter and collector resistances are higher for the *pnp* over the *npn* device. Although for simplicity we have assumed complete activation of the dopant concentration, for fabricated devices, solid-solubility limits will be another constraint that will induce differences in the emitter, base and collector contact resistances of these devices. In general, for fabricated devices with higher contact resistances, parasitics will play a more dominant role in lowering the f_{MAX}.

For identical doping concentrations in all regions of the *npn* and *pnp* devices, the *npn* provides superior performance compared to the *pnp*, as expected. To achieve optimized complementary SiGe HBT profiles with comparable performance for any target technology node, the best *npn* performance is generally reduced to match the performance of the optimized *pnp* HBT. Once the dopant profiles are inverted from the *npn* to the *pnp* devices at the same concentration, the doping and Ge profiles at the CB and EB junction are key

elements which require redesign and fine tuning. *ac* simulations were performed for a wide range of V_{CB} and V_{CE} values to optimize the candidate devices towards matched performance for important biasing topologies used in circuit applications. This is important to ensure that the devices do not suffer from any unoptimized HBE when driven into saturation by a low supply voltage, in low-power applications which simultaneously require high-performance.

3.3 200 GHz Complementary SiGe HBT Device Optimization Results

A comparison of the final optimized candidate *npn* and *pnp* device profiles for a 200 GHz C-SiGe HBT technology at the 120 nm scaling node is shown in Fig. 3. While the Ge profiles are not significantly different, the *pnp* requires a much larger SIC doping to achieve a comparable performance and delay the onset of HBE, even with a similar Ge retrograde to the *npn*. The *dc* and *ac* performance of the 200 GHz candidate device profiles are shown in Figs. 4 and 5. While the maximum β_{DC} (though it occurs at a higher J_C for the *pnp*), I_C at same I_B, and peak f_T are quite comparable for the C-SiGe devices, peak f_{MAX} for the *npn* is higher than for the *pnp* due to a lower SIC doping. Although the f_T and f_{MAX} values are higher for the *npn* at any J_C below the peak values, the peak f_T and f_{MAX} occurs at slightly higher J_C for the *pnp* device due to a higher SIC doping, as the onset of both Kirk-effect and the HBE that causes the f_T/f_{MAX} roll-off is delayed to higher J_C in the *pnp*. BV_{CEO} extracted at moderate injection using the base-current reversal point under forced-V_{BE} conditions are 1.78 V and 1.97 V for the *npn* and the *pnp* devices, respectively. This is mainly due to a lower M-1 for the *pnp* compared to the *npn* even with a higher collector doping. At similar J_C and V_{CB} values, the *pnp* device will show slightly higher self-heating (J^*E) over the *npn* due to a larger CB junction electric field (E) resulting from higher doping.

Although these device profiles were initially matched for performance under a single bias condition, the profiles were further optimized for comparable performance over a range of V_{CE} and V_{CB} values. As shown in Fig. 6, the optimized 200 GHz complementary HBT

profiles compare very well in their performance over a wide bias range. The regional T-T analyses of the optimized *npn* and *pnp* profiles are shown in Fig. 7. For both devices, the EB junction T-T (τ_{be}) limits performance at low-to-moderate injection, while CB junction T-T ($\tau_{bc}+\tau_c$) limits the performance at moderate-to-high injection (at or around peak f_T). The base T-T (τ_b) limits the performance in the very high-injection regime of device operation (well beyond peak f_T). Even at a comparable doping, the *pnp* will have a higher R_e, R_b and R_c than the *npn*. This contributes to a lower f_T and f_{MAX} for the *pnp* at J_C below peak f_T [1].

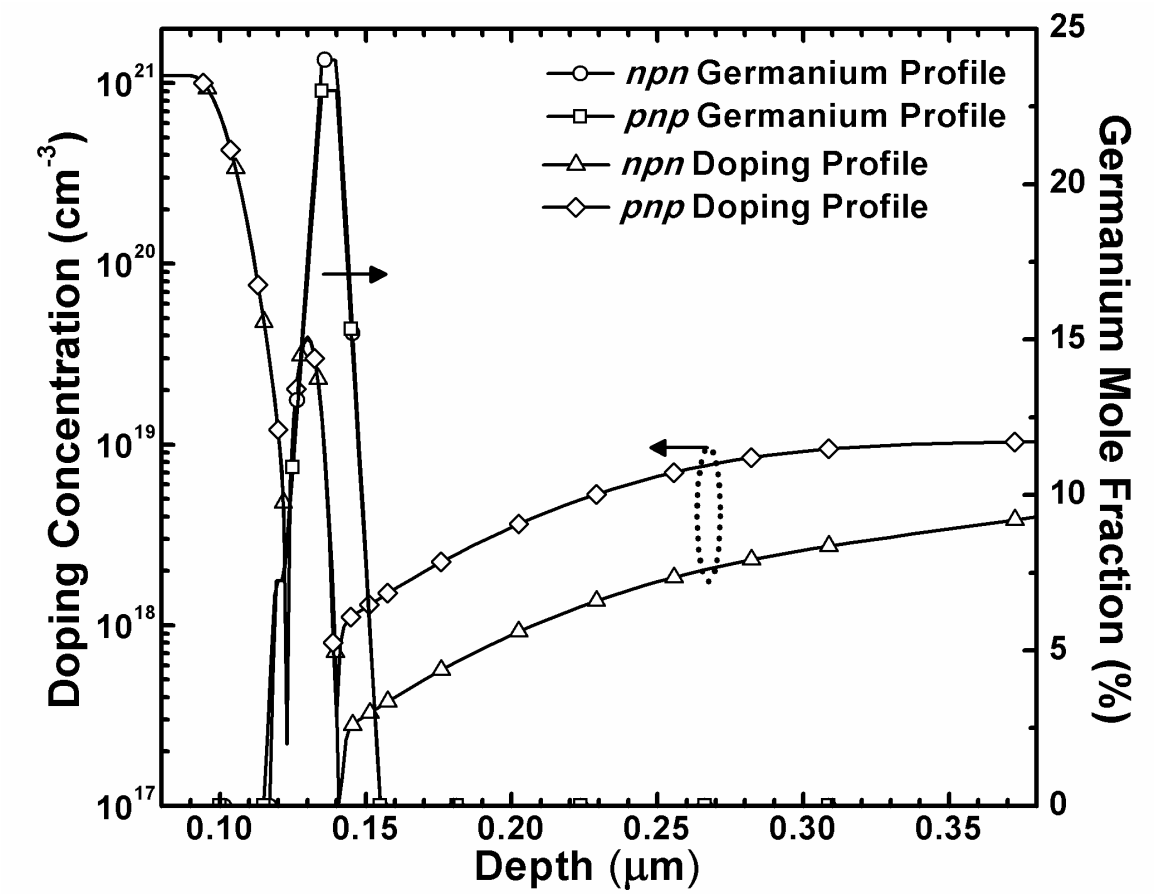

Figure 3: Doping and Ge profiles for optimized complementary SiGe HBT device structures with 200 GHz performance [2].

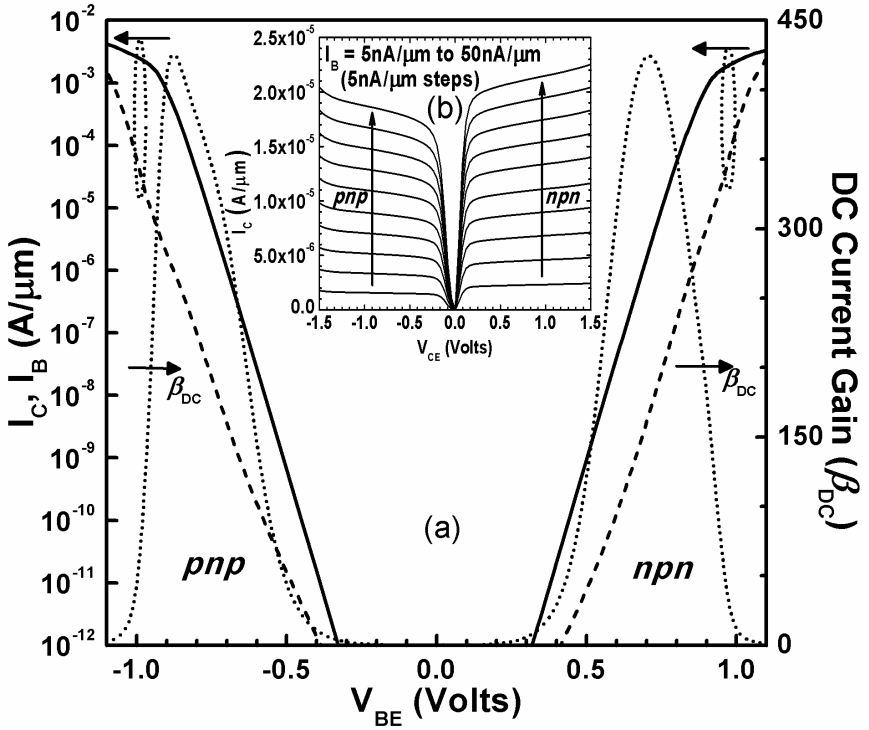

Figure 4: Comparison of the (a) dc performance (Gummel plots and current gain), and the (b) output characteristics for the optimized 200 GHz *npn* and *pnp* HBT device profiles in Fig. 3 [2].

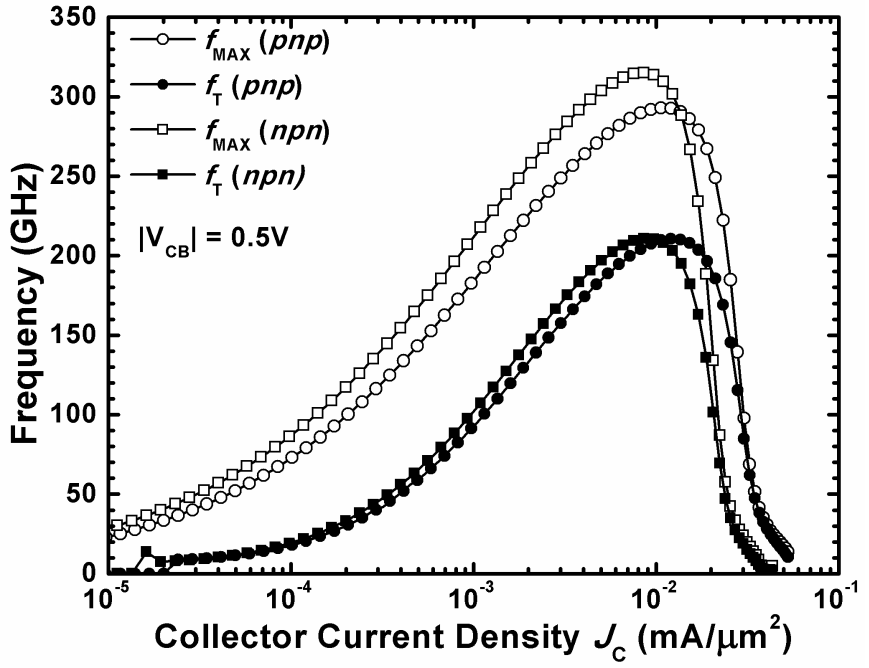

Figure 5: Comparison of the ac performance (f_T, f_{MAX}) for the optimized 200 GHz *npn* and *pnp* HBT device profiles in Fig. 3 [2].

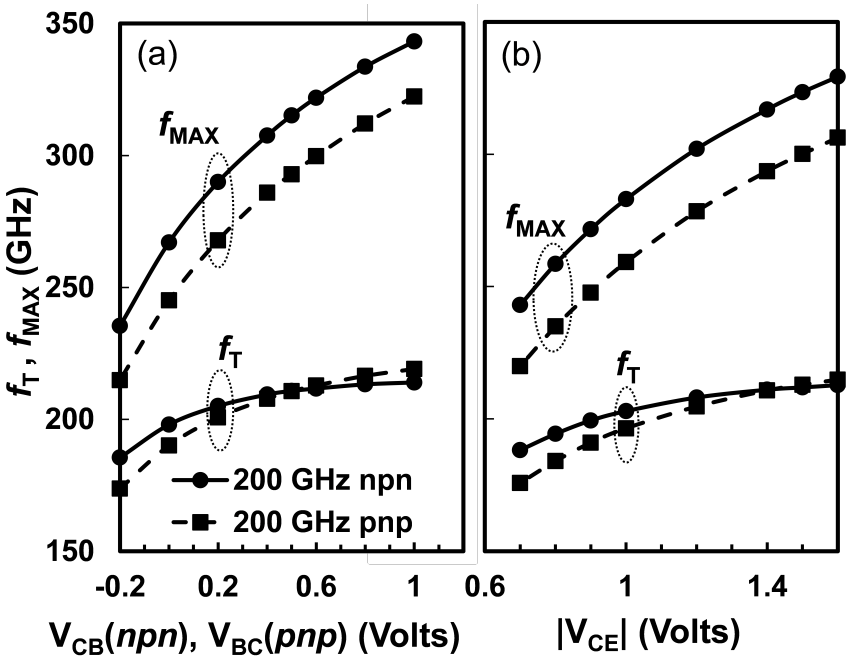

Figure 6: Comparison of the simulated peak f_T and f_{MAX} for the optimized 200 GHz *npn* (and *pnp*) HBT profiles in Fig. 3 obtained at different (a) V_{CB} (V_{BC}) and (b) $|V_{CE}|$ values [2].

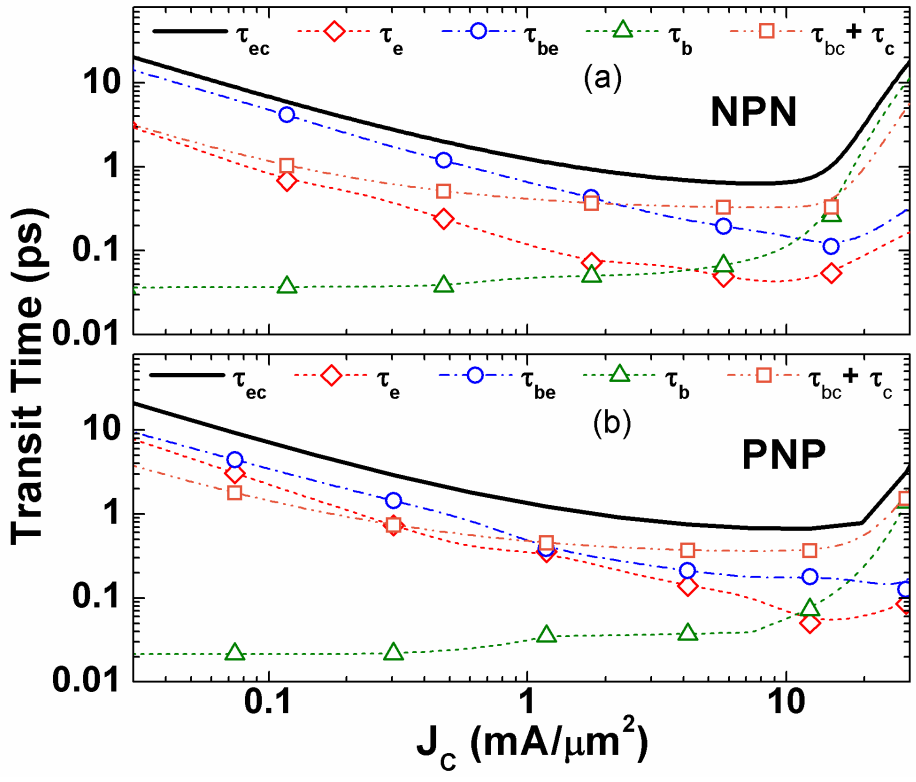

Figure 7: Regional transit-time parameters extracted from regional quasi-static transit-time analysis of the 200 GHz (a) *npn* and (b) *pnp* device profiles in Fig. 3 at $|V_{CB}|$ = 0.5V [2].

3.4 Complementary SiGe HBT Performance Scaling Roadmap

To demonstrate the utility of the developed integrated simulation framework towards developing a C-SiGe HBT scaling roadmap, C-SiGe devices were also developed for a target of 100 GHz peak f_T at the 180 nm lithography node. The optimized *npn* and *pnp* device profiles for the 100 GHz node are shown in Fig. 8. The 100 GHz profiles are in some ways similar to those of the 200 GHz profiles. The *pnp* has a higher SIC doping, slightly higher Ge content in the base, and a comparable retrograde. The *ac* performance of the 100 GHz *npn* and *pnp* devices compare very well over a wide range of V_{CB} and V_{CE} values as shown in Fig. 9.

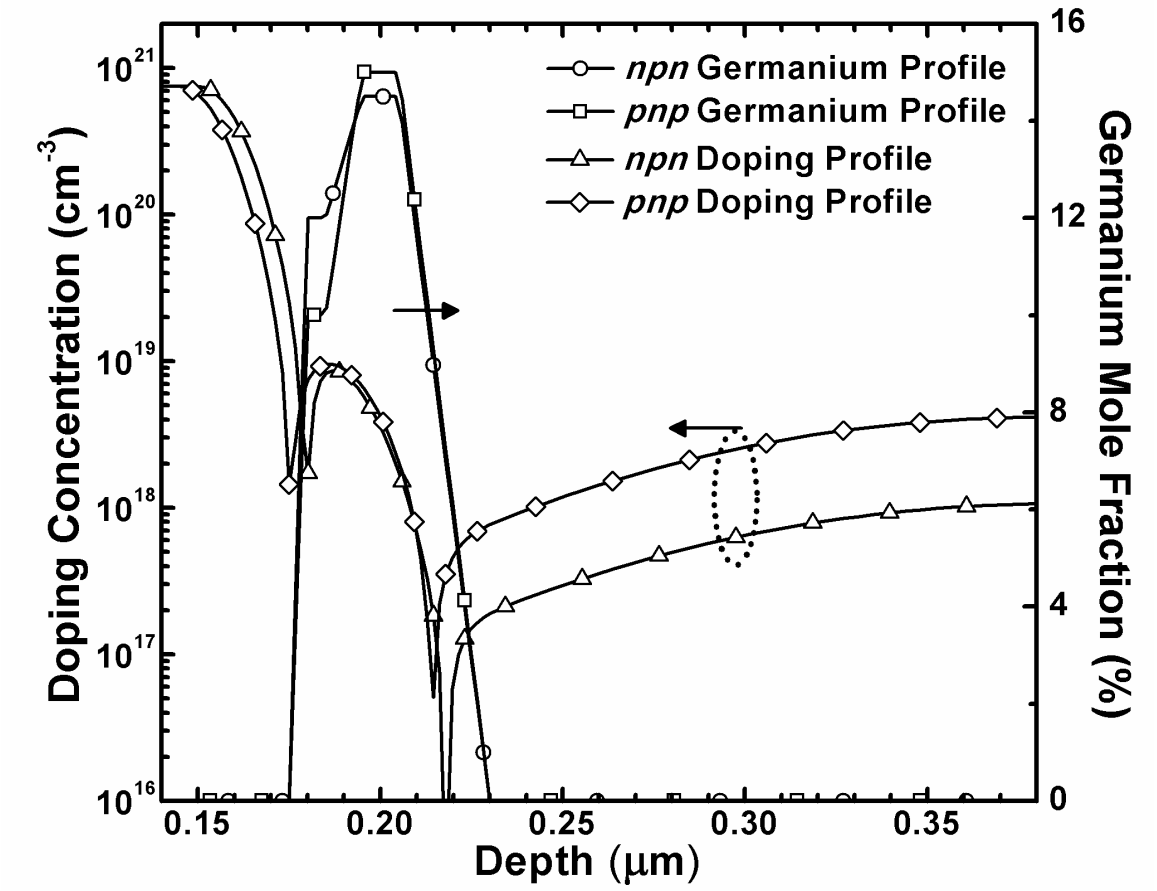

Figure 8: Doping and Ge profiles for optimized complementary SiGe HBT device structures with 100 GHz performance [2].

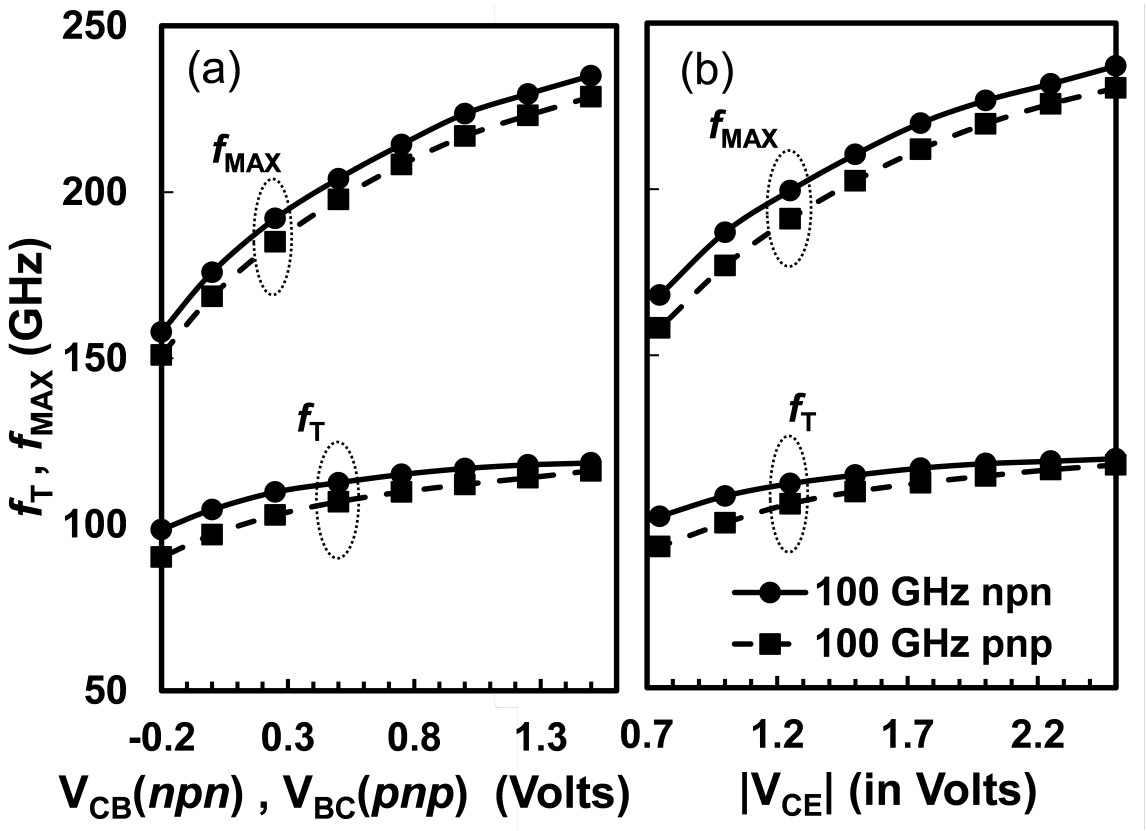

Figure 9: Comparison of the simulated peak f_T and f_{MAX} for the optimized 100 GHz *npn* (and *pnp*) HBT profiles in Fig. 8 obtained at different (a) V_{CB} (V_{BC}) and (b) $|V_{CE}|$ values [2].

Table 1 shows a list of some of the important 2-D physical parameters (given in Fig. 1) at each performance node. The parameters were kept identical for the *npn* and *pnp* devices. Between the 100 GHz and 200 GHz devices, the only physical parameter varied was emitter width W_E through the variable EW. Table 2 shows a summary of all the important FoM for the C-SiGe devices at both the 100 and 200 GHz nodes, displayed as a performance roadmap. Fig. 10 shows regional T-T components for each of these devices as a fraction of the total emitter-to-collector transit time (τ_{ec}) at the peak f_T condition. The performance roadmap in Table 2 can be extended using the same simulation framework to include other performance nodes at $\leq$ 200 GHz.

Table 1: Important 2-D Physical Parameters for the Optimized Device Structures [2].

Performance Node	200 GHz	100 GHz
EW (W_E (nm) = 2EW)	60 nm	90 nm
EBOD	0	0
EBOSEP	60 nm	60 nm
EV1	55 nm	55 nm
Emitter-to-STI Spacing (STIS)	290 nm	290 nm
STI Depth (= STIV2 – STIV1)	340 nm	340 nm
STI Width (STIW)	180 nm	180 nm
STI-to-DT Spacing (CW)	300 nm	300 nm
STIV1	160 nm	160 nm
BW	410 nm	410 nm

Table 2: Simulated Figures-of-Merit for the Complementary SiGe Devices as part of the Scaling Roadmap [2].

Performance Node and Device Type	200 GHz *pnp*	200 GHz *npn*	100 GHz *pnp*	100 GHz *npn*
Emitter Width W_E (nm)	120	120	180	180
Maximum β_{DC}	424	421	232	200
R_{BI} (kΩ/□)	4.77	2.91	8.12	6.91
BV_{CEO} (V)	1.97	1.78	2.82	2.65
Peak f_T (GHz)	211	211	107	113
J_c@ Peak f_T (mA/μm^2)	12.1	8.6	4.3	4.1
Peak f_{MAX} (GHz)	293	315	198	204
J_c@ Peak f_{MAX} (mA/μm^2)	10.7	8.6	3.3	3.08

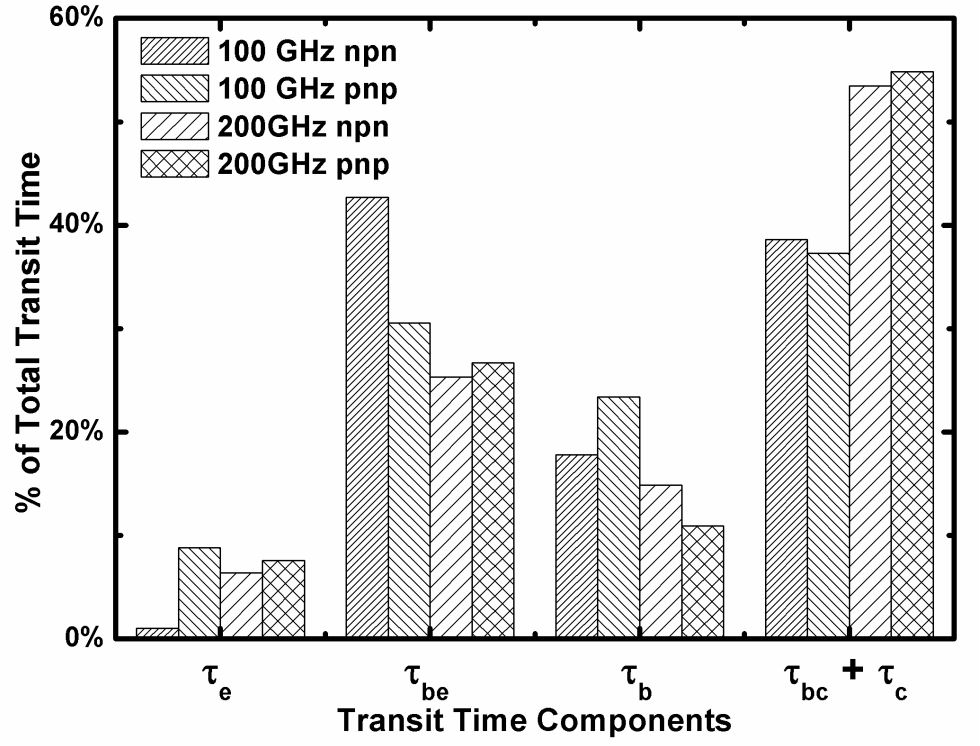

Figure 10: Simulated regional transit-time components at the peak f_T condition (at $|V_{CB}|$ = 0.5 V) as a fraction of the total transit time for all the optimized 100 GHz and 200 GHz *npn* and *pnp* HBT profiles (in Figs. 3 and 8) [2].

It is equally important to take into consideration additional reliability FoM during the TCAD-based device performance optimization process; this has never previously been attempted in a TCAD environment for device optimization. For example, the reverse-biased (RB) current or the zero-bias peak electric field at the EB junction holds a direct correlation to the long-term reliability of the device. This is particularly important for optimization of C-SiGe HBTs that need to be comparable in both their performance and reliability. A recent investigation using the integrated simulation framework from this study showed that the optimized C-SiGe HBT profiles developed here compared very well in the simulated reverse-biased EB junction tunneling current at both the 100 and 200 GHz performance nodes and will be discussed here in Section 3.4 [5]. In this context, it is important to perform simultaneous predictive estimation of reliability within a device optimization framework used in this study. While it is important to achieve comparable performance and reliability when scaling C-SiGe HBTs, it is equally pertinent to explore and develop new

applications that will utilize greater device performance, while pushing the performance of existing C-SiGe circuits [67, 68].

When the performances of the device profiles optimized here are compared to that of existing C-SiGe HBTs reported in the literature on a f_T vs. BV_{CEO} plot in Fig. 11, the f_T vs. BV_{CEO} scaling tradeoff commonly known as the Johnson limit is clearly observed. The 100 GHz optimized devices are very comparable to those reported in the literature with similar $f_T \times BV_{CEO}$ product, further indicating the robustness and predictive nature of the device optimization methodology implemented in this study. With the increase of f_T resulting from device scaling, the C-SiGe devices continually move to a higher $f_T \times BV_{CEO}$ product (dotted lines) due to an increase in the SIC doping with performance, similar to that reported for scaling of *npn* devices in earlier studies [62, 69]. This demonstrates the feasibility of a performance scaling roadmap for C-SiGe HBTs akin to that for *npn* SiGe HBTs [62].

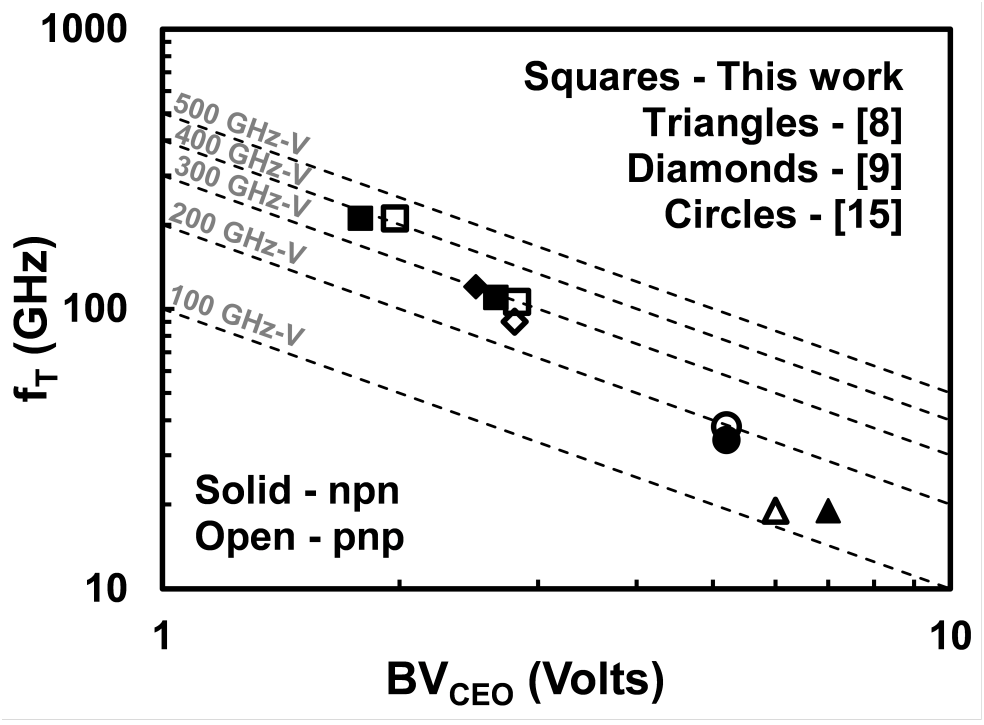

Figure 11: f_T vs. BV_{CEO} plot of the optimized complementary SiGe HBT devices from this study, compared to actual devices reported in the marked citations within [2], showing the tradeoff between performance and operating voltage. [2].

If the performances of C-SiGe devices from this study are compared to those reported in the literature on a f_T vs. f_{MAX} plot as in Fig. 12, the devices can be clearly grouped into three generations (peak f_T of ≤50, ~100, and ~200 GHz with typically $f_{MAX} > f_T$ for each generation). This is comparable to the three generations of existing *npn*-only SiGe HBTs (based on the constant $f_T + f_{MAX}$ lines), indicating that performance scaling for the C-SiGe HBTs can be enabled through successful fabrication of the C-SiGe devices.

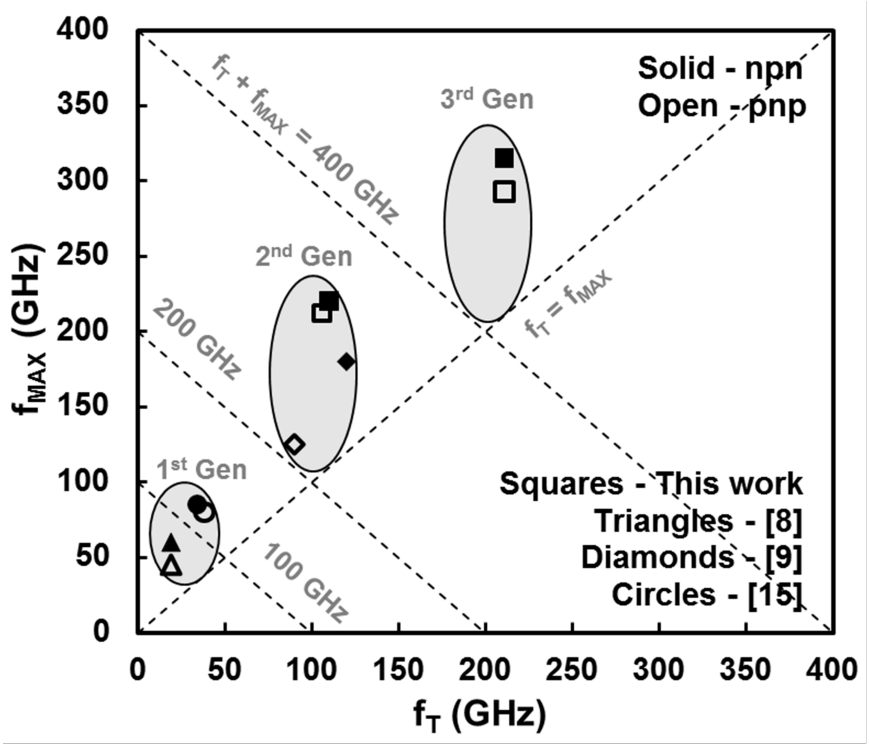

Figure 12: Peak f_{MAX} vs. f_T plot of the optimized complementary SiGe HBT devices from this study, compared to actual devices reported in the marked citations within [2], showing distinct generations of complementary SiGe device performance with $f_{MAX} > f_T$ [2].

3.5 Summary

This study has for the first time successfully developed an integrated TCAD simulation framework and methodology for predictive optimization and scaling of C-SiGe HBTs to achieve comparable performance and reliability. The utility of this framework has been demonstrated by showing for the first time a path towards development of a performance scaling roadmap for C-SiGe HBTs. This integrated simulation framework lends itself to

post-processing and analysis and is highly flexible for use in any kind of technology development environment within the semiconductor industry.

Within the scope of the simulation methodology highlighted here, this study also proves for the first time that performance optimization and scaling of C-SiGe HBTs is feasible, just as in *npn*-only technologies, as long as there are existing methods to commercially fabricate these devices. Any inaccuracy in the *pnp* transport model in the absence of calibration to fabricated devices will clearly induce errors in f_T, f_{MAX}, BV_{CEO}, etc. However, this would only require some additional fine tuning of the current candidate *pnp* device profiles (mainly doping in the base and collector, and the Ge profile) to achieve performance comparable to the *npn*. This study also demonstrates for the first time an integrated method to optimize C-SiGe HBTs for comparable performance over a wide range of bias values, important for optimizing devices with a range of low-power and high-performance applications. Considering that complementary bipolar technologies will remain very attractive for high performance, high frequency, analog, and mixed-signal circuits, scaling and development of such technologies up to 200 GHz performance and beyond should provide major breakthroughs in that application space.

CHAPTER 4

ELECTRO-THERMAL CONSTRAINTS FOR SCALING SIGE HBTS ON THICK-FILM SOI

In this section, npn SiGe HBTs from a commercially-available 150 GHz bulk-Si BiCMOS process [70] with peak f_T/f_{MAX} of 150/180 GHz were fabricated on a thick-film SOI substrate and compared with the bulk control. This investigation explores the self-heating effects on the *dc* and *ac* performance scaling of the SOI and bulk devices resulting from electro-thermal feedback, and introduces novel characterization methods to study these devices [3].

The devices in the present investigation were studied for self-heating effects using different biasing techniques in order to characterize their behavior around the onset of strong electro-thermal (ET) feedback, as well as to explore the consequent electrical constraints imposed on the device operating conditions due to thermal runaway (TR). Specifically, the effects on device linearity and RF SOA was studied and implications were analyzed (to be submitted for publication).

4.1 Experimental Details

The cross-section for the bulk-Si device with STI and DT can be found in [70]. The devices used in this study were fabricated with the same doping and Ge profiles and an identical process flow, with the exception that they were placed on a thick-film SOI substrate instead of a bulk-Si substrate in the control.

dc measurements were performed using an Agilent 4156B Parameter Analyzer. *ac* measurements on these devices were performed using an Agilent E8361C PNA. Pulsed measurements were performed with a DIVA D210 dynamic I-V analyzer from Accent Optical Technologies.

The devices were biased in forced-V_{BE} (FVB) and forced-I_B (FIB) modes for common-emitter operation. Different single-finger emitter geometries were investigated, and the results shown are typical for the different devices sizes. All measurements were performed at room temperature. For the purpose of this discussion, the term "thermal runaway (TR)" is not used for a catastrophic failure within the device, but to indicate an unstable region of device operation with positive electro-thermal feedback.

4.2 Discussion of *dc* and *ac* Measurement Results

Since the device profiles were optimized for bulk-Si substrates, but were instead fabricated on an SOI substrate for the purposes of this study, differences between the bulk and SOI devices will be highlighted here. The BV_{CEO} of the bulk device is reported to be 2.3 V [70]. Fig. 13 show a comparison of the FVB Gummel characteristics for the SOI and bulk devices, and their evolution with increasing V_{CB}. The results on the SOI device show increased I_C and I_B at high injection even at V_{CB} = 0 V (minimal self-heating compared to bulk). The high-injection I_C and I_B show a significant increase with V_{CB} for the SOI device compared to bulk, indicating stronger self-heating and electro-thermal feedback, as expected. Fig. 14 show the corresponding normalized *dc* current gain (β_{DC}) for the two devices and its evolution with increasing V_{CB}. It is clearly evident that the SOI device shows a sharper collapse in β_{DC} at high injection, and this can be correlated to a greater proportional increase in I_B compared to I_C in Fig. 13 after the onset of strong electro-thermal feedback (around 0.87 V). The increase in I_C and I_B with self-heating, and the corresponding collapse in β_{DC} becomes sharper at higher V_{CB} values. Similar evolution of the Gummel characteristics was also observed under FVB measurements with constant V_{CE}. Onset of TR in these plots can be characterized as the point where $\partial I_C/\partial V_{BE} \rightarrow \infty$. This threshold of ET instabilities moves to lower V_{BE} with increasing V_{CE} (or V_{CB}).

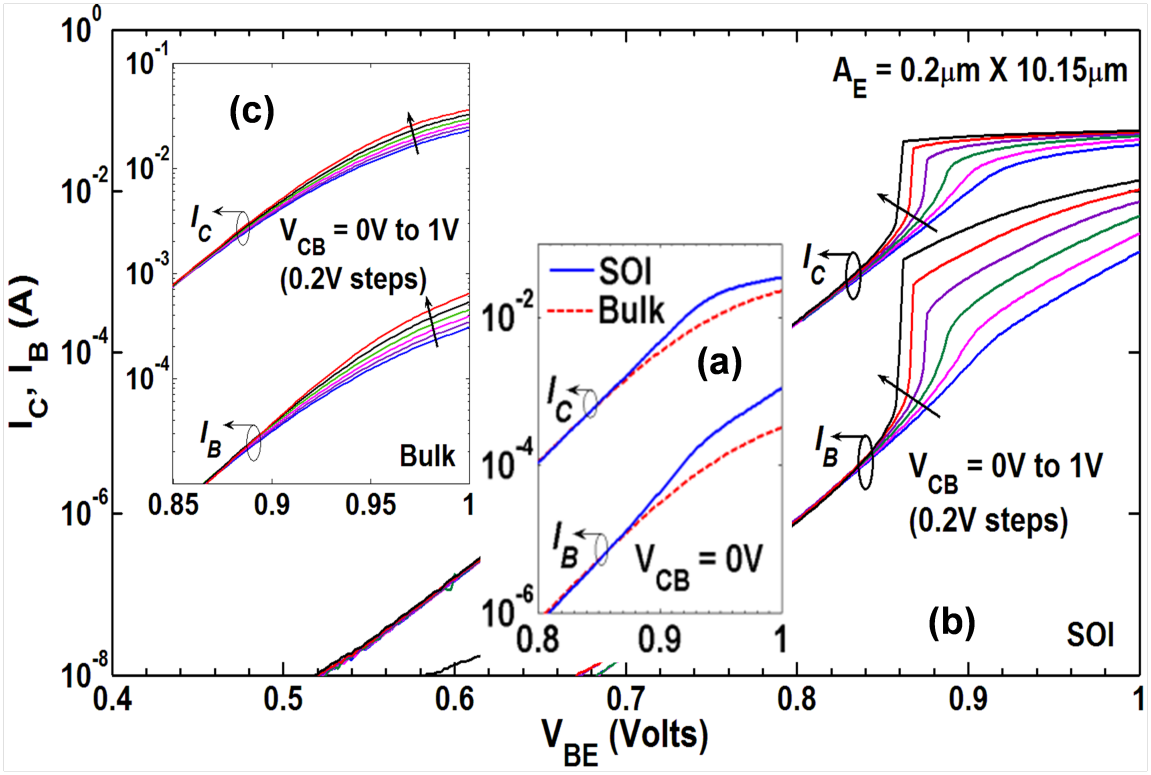

Figure 13: (a) Gummel characteristics for the SOI and Bulk devices at V_{CB} = 0 V. Evolution of Gummel plots with increasing V_{CB} for the (b) SOI and (c) Bulk device. All measurements were under forced-V_{BE} condition [3].

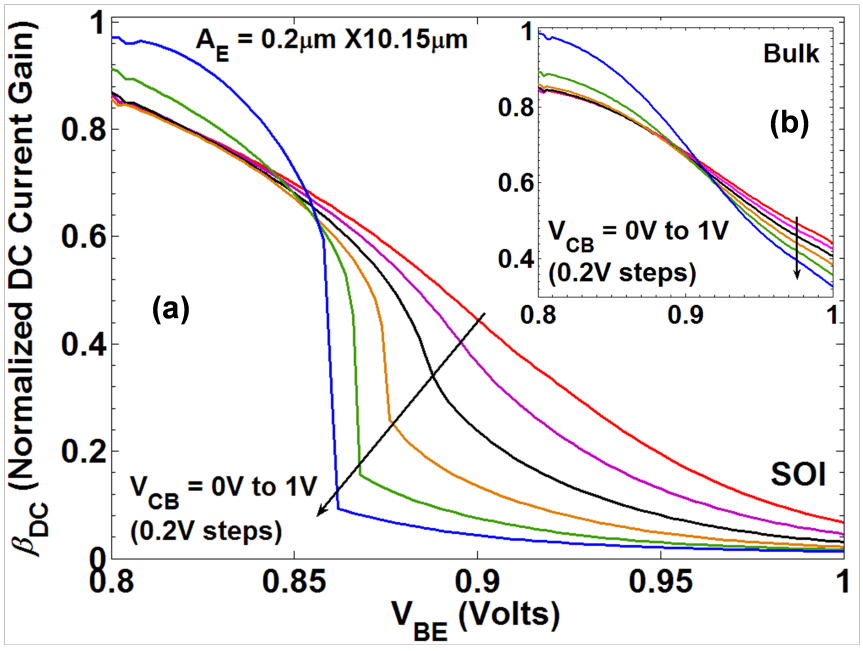

Figure 14: Evolution of the DC Current Gain with increasing V_{CB} under forced-V_{BE} measurement for the (a) SOI and (b) Bulk device [3].

However when the Gummels for the SOI device were measured with a FIB method at a constant V_{CE}, both I_C and I_B shows a negative differential resistance (NDR or $\partial V_{BE}/\partial I_C < 0$) region or snap-back beyond the onset of TR ($\partial I_C/\partial V_{BE} \rightarrow \infty$), as is evident from Fig. 15. Analytical formulations for explaining the differences observed between the current and voltage controlled measurements based on ET feedback in the device have been discussed in [39, 71]. These observed differences are further confirmed in Fig. 16, where FVB output characteristics (OC) show a sharp increase in I_C beyond the onset of TR (observed at $\partial I_C/\partial V_{CE} \rightarrow \infty$), while the FIB measurement shows a NDR region ($\partial V_{CE}/\partial I_C < 0$) beyond the $\partial I_C/\partial V_{CE} = 0$ point caused by current gain collapse due to SH. No significant differences were observed between FVB and FIB measurements of the Gummel and OC for the SOI device at low injection; and for the bulk control device over all bias levels due to minimal SH in both devices under these conditions.

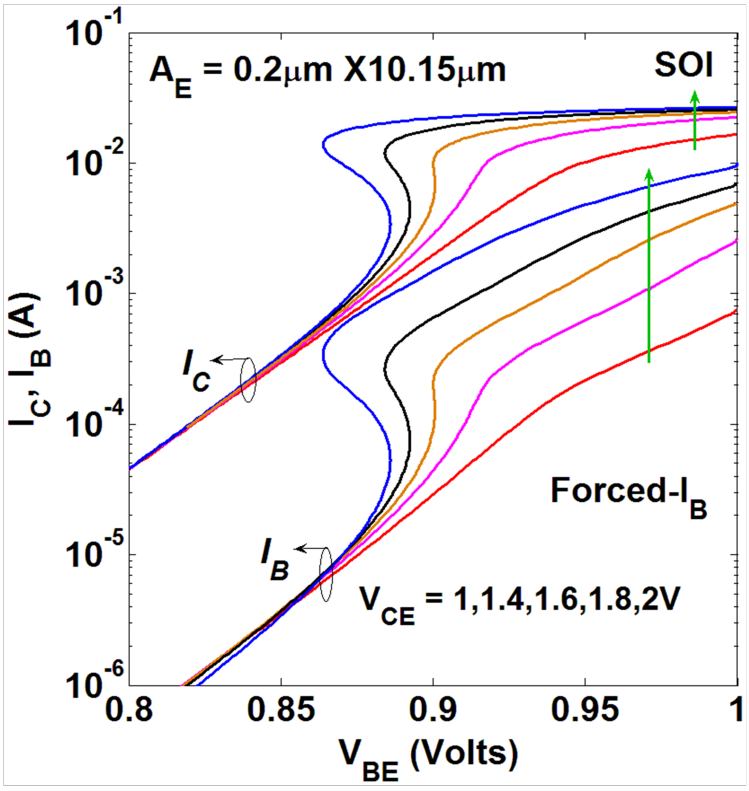

Figure 15: Evolution of the Gummel characteristics of the SOI device with increasing V_{CE} under forced-I_B measurement [3].

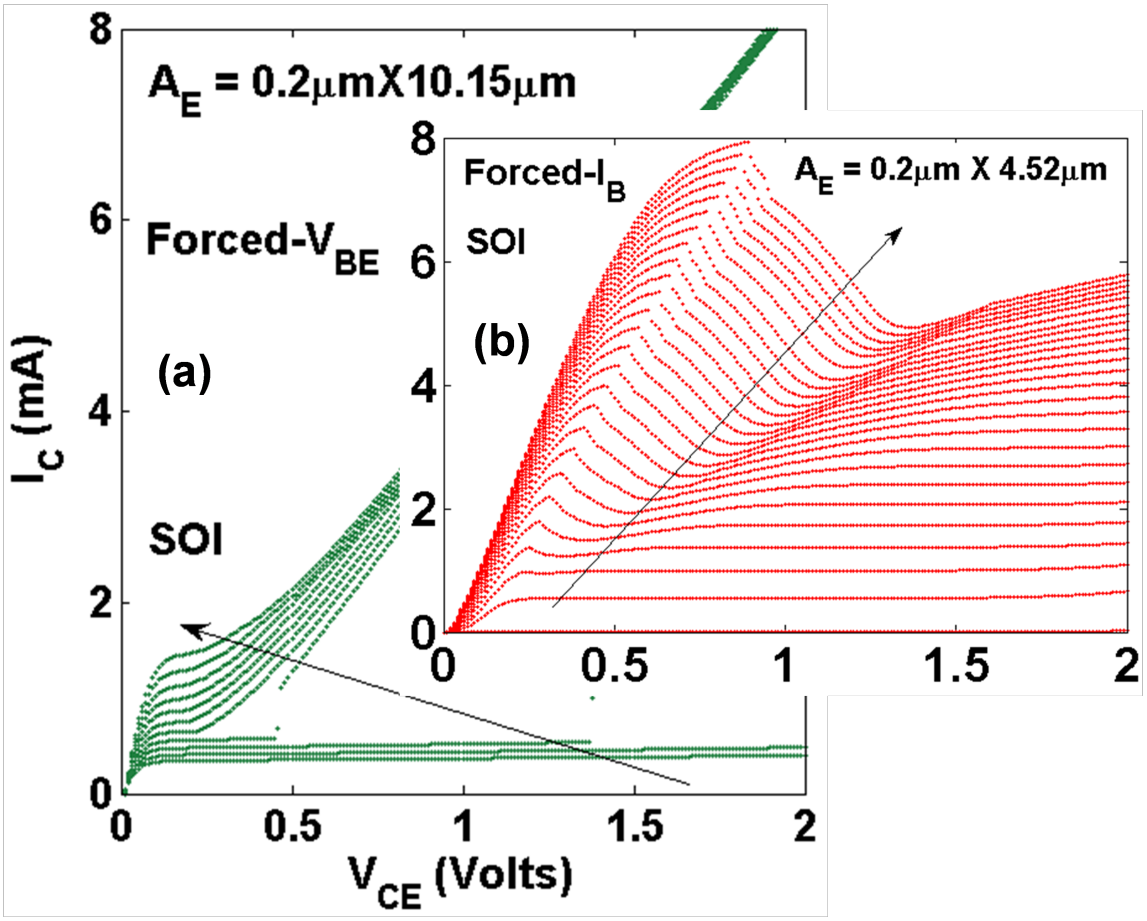

Figure 16: DC output characteristics of the SOI device measured under (a) forced-V_{BE}, and (b) forced-I_B conditions [3].

FVB *ac* measurements on the SOI and bulk devices are compared in Fig. 17 for $V_{CB} = 0$ V (minimal SH condition). Since at high-injection the SOI device has higher I_C at the same V_{BE} due to SH (as shown in Fig. 13), the f_T/f_{MAX} roll-off due to Kirk-effect starts at a lower V_{BE}, and thus the peak f_T moves to a lower V_{BE} for the SOI device in Fig. 17(a). However, from the f_T/f_{MAX} vs. I_C plot in Fig. 17(b) it is evident that the roll-off from peak f_T/f_{MAX} still happens at a comparable I_C (or J_C) for both devices, thereby implying that even with the increase in I_C due to higher SH in the SOI device, the f_T/f_{MAX} roll-off is fundamentally driven by onset of the Kirk-effect mechanism resulting from the increase in I_C.

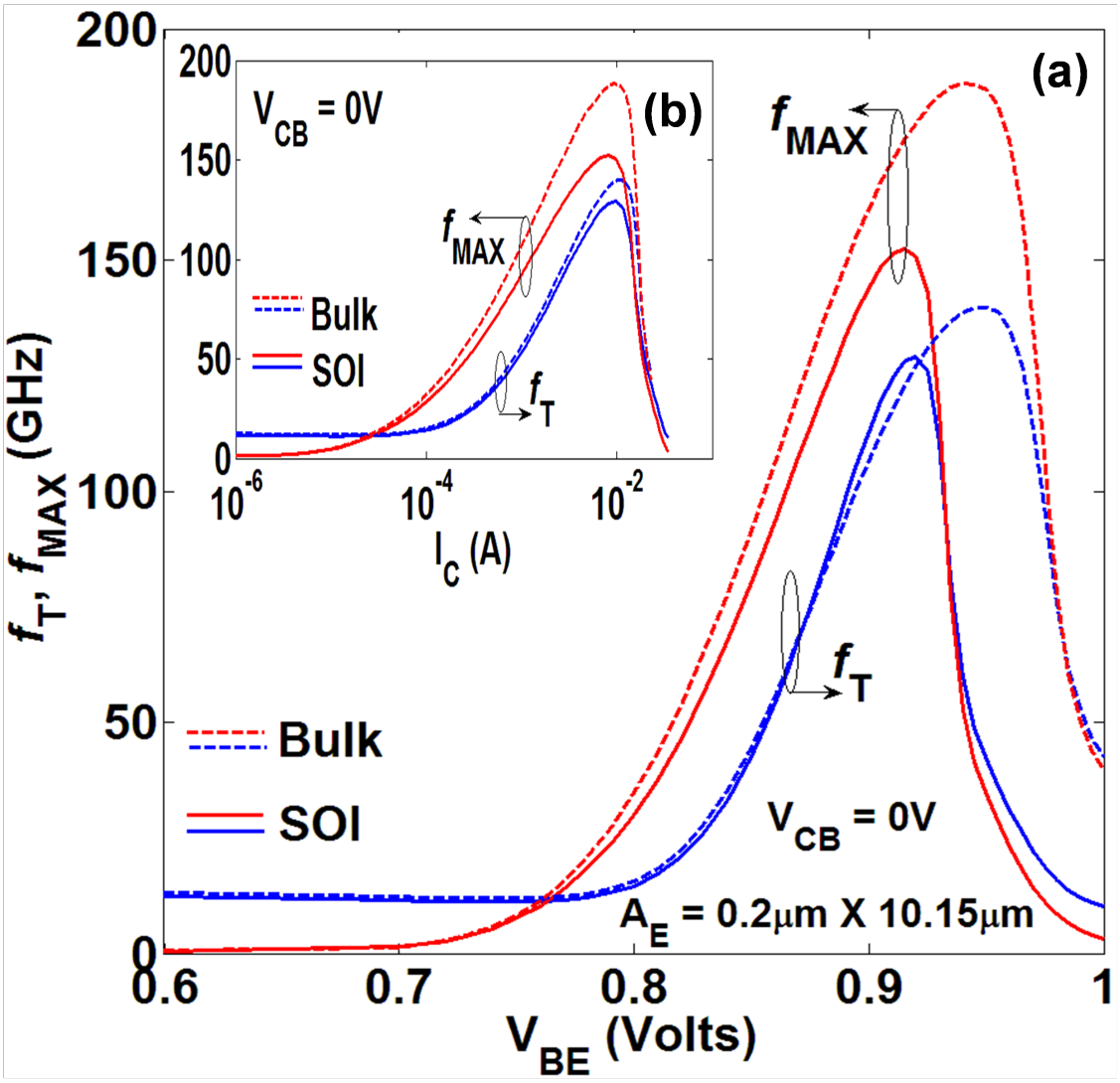

Figure 17: Comparison of f_T, f_{MAX} for the SOI and the bulk devices measured at V_{CB} = 0V under forced-V_{BE} conditions, plotted versus (a) V_{BE}, and (b) I_C [3].

The J_C at the onset of Kirk-effect is analytically given by the following expression [8].

$$J_{C,Kirk} = qv_s N_{DC}\left\{1+\frac{2\varepsilon(V_{CB}+\phi_{bi})}{qN_{DC}W_{epi}^2}\right\} \tag{1}$$

It can be inferred from the above expression that ϕ_{bi} (built-in potential) and v_s (saturation drift velocity) are the prime temperature sensitive parameters that determine $J_{C,Kirk}$, as a function of the relevant device temperatures. It is known that v_s decreases with increasing temperature [72]; and ϕ_{bi} also decreases weakly with temperature [73].

FVB f_T/f_{MAX} plots of the SOI device measured at different V_{CB} are shown in Fig. 18, and shows that the onset of f_T/f_{MAX} roll-off moves to lower V_{BE} with increasing V_{CB} (or V_{CE}). This behavior is primarily determined by the onset of strong ET feedback and consequent I_C increase in the device. Thus, SH effectively limits usable V_{BE} (to a critical value we term $V_{BE,crit}$) at a given V_{CB} (or V_{CE}) before the onset of strong positive ET feedback leads to a sharp increase in I_C and consequent Kirk effect driven f_T/f_{MAX} roll-off. In other words, there will be a maximum V_{CE} (say $V_{CE,crit}$) which can be applied to the device with a specified I_C (say at peak f_T condition) before the onset of TR. The analytical treatment for this argument has been reported in [37]. The $V_{BE,crit}$ values extracted for the SOI device under FVB method of operation are plotted against applied V_{CB} (or V_{CE}) for two different device sizes in Fig. 19.

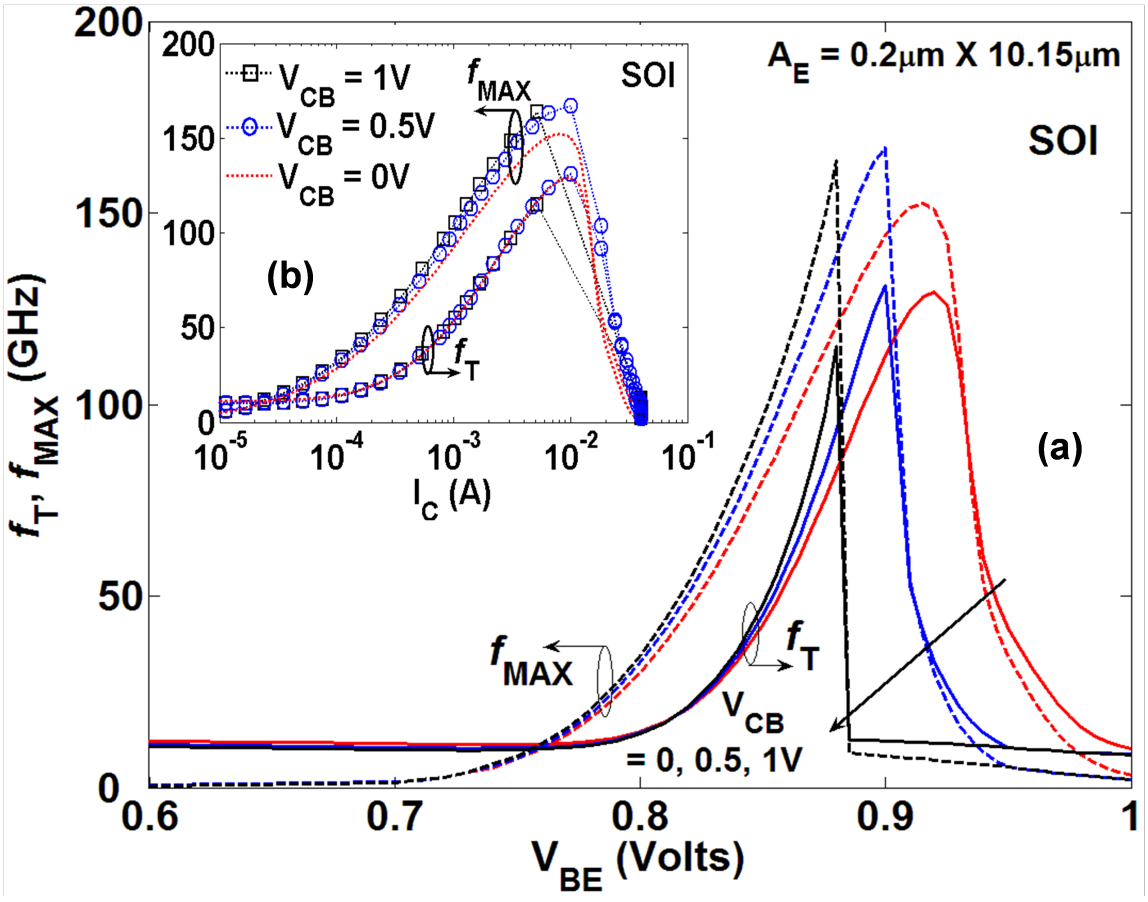

Figure 18: Comparison of f_T, f_{MAX} plots for the SOI device measured at different V_{CB} under forced-V_{BE} conditions, plotted versus (a) V_{BE}, and (b) I_C [3].

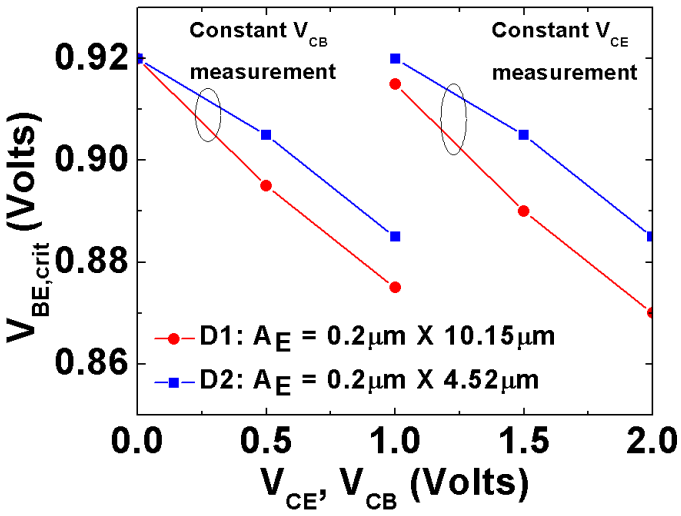

Figure 19: Plot of the $V_{BE,crit}$ FOM vs. applied V_{CB} and V_{CE} for two different sizes of the SOI device showing similar trends [3].

However, when f_T/f_{MAX} is plotted vs. I_C in Fig. 18, the peak f_T/f_{MAX} remains at a comparable I_C with increasing V_{CB} (and SH); until at V_{CB} = 1 V (with high enough SH), when strong positive ET feedback leads to TR, causing a sudden increase in I_C and a very sharp resultant f_T/f_{MAX} roll-off (within ~1-2 mV) due to Kirk-effect. Sudden change in I_C prevents biasing the device with finer I_C intervals after the onset of TR, leading to scarcity of measurement points and giving rise to a triangular shape when individual points are connected. This proves that $J_{C,Kirk}$ variations are weak for the relevant device temperatures resulting from SH within these devices. No similar observable SH effects were noted in the bulk device under identical measurement conditions.

Next, for the first time, an attempt is made to measure f_T/f_{MAX} of the SOI device under FIB biasing conditions to gain further insight into the *ac* performance of the device in the NDR region of operation beyond the onset of TR, as the sudden f_T/f_{MAX} roll-off observed beyond TR under FVB condition failed to provide this information. Fig. 20 compares f_T/f_{MAX} vs. I_C measured at constant V_{CE}, with FVB and FIB methods for both SOI and bulk devices. As can be inferred from the plots for V_{CE} = 1 V (with minimal SH), that there is no observable difference in the measured f_T/f_{MAX} between the two measurement methods. With increased SH and ET feedback at V_{CE} = 2 V, the FIB method is able to

measure f_T/f_{MAX} of the SOI device after the onset of thermal runaway ($\partial I_C/\partial V_{BE} \rightarrow \infty$) up to I_C close to peak f_T/f_{MAX} and then into the roll-off region with more details than the FVB method. Under current-controlled (FIB) operation, positive ET feedback and hence SH of the device is limited, which consequently prevents any sudden change in I_C leading to TR, thereby allowing *ac* measurements at finer I_C intervals up to and beyond peak f_T/f_{MAX}, which is not possible with the FVB method. Beyond peak f_T/f_{MAX}, the NDR region shows up as a unique concave region on the f_T/f_{MAX} vs. I_C plot (as shown in Fig. 20(b)), followed by a typical roll-off due to Kirk effect. This further proves that the I_C at peak f_T/f_{MAX} is fairly stable, and is determined by a stable $J_{C,Kirk}$ (which is a weak function of temperature) even under strong SH effects in these devices. There is no observed difference in the f_T/f_{MAX} of the bulk device from the two methods (as shown in Fig. 20(b)).

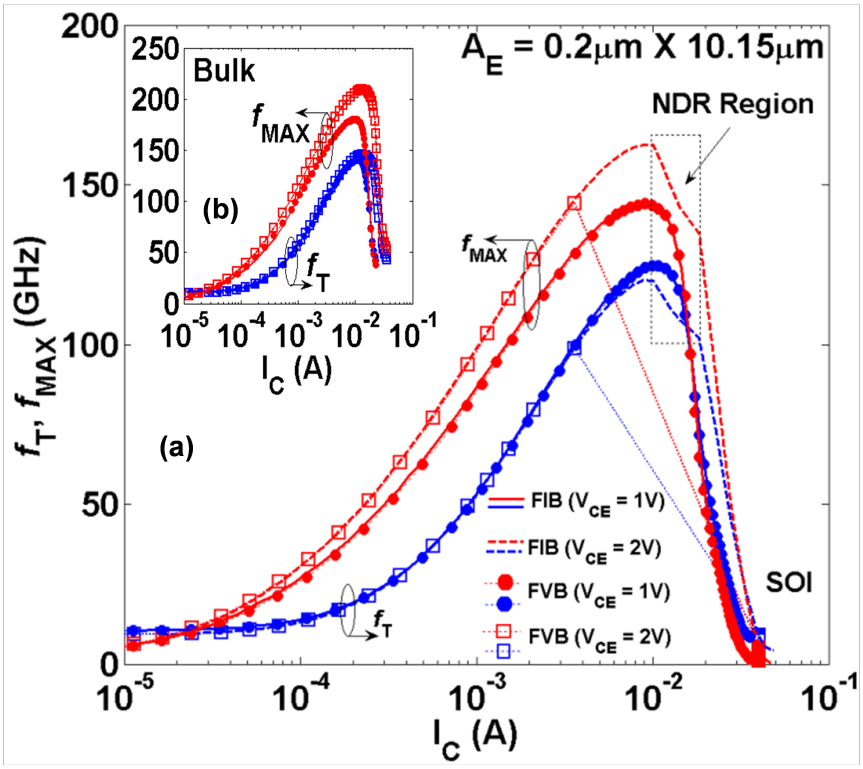

Figure 20: (a) Comparison of f_T, f_{MAX} for the SOI device measured at different V_{CE} under forced-V_{BE} (FVB) and forced-I_B (FIB) conditions. (b) Plots for the bulk device measured under similar conditions [3].

FVB mode OC from pulsed $I - V$ measurements on the SOI and bulk devices are shown in Fig. 21. Although the measurement setup is limited to the high-injection regime, stronger SH effects are clearly evident for the SOI device under *dc* conditions, leading to a greater difference in I_C.

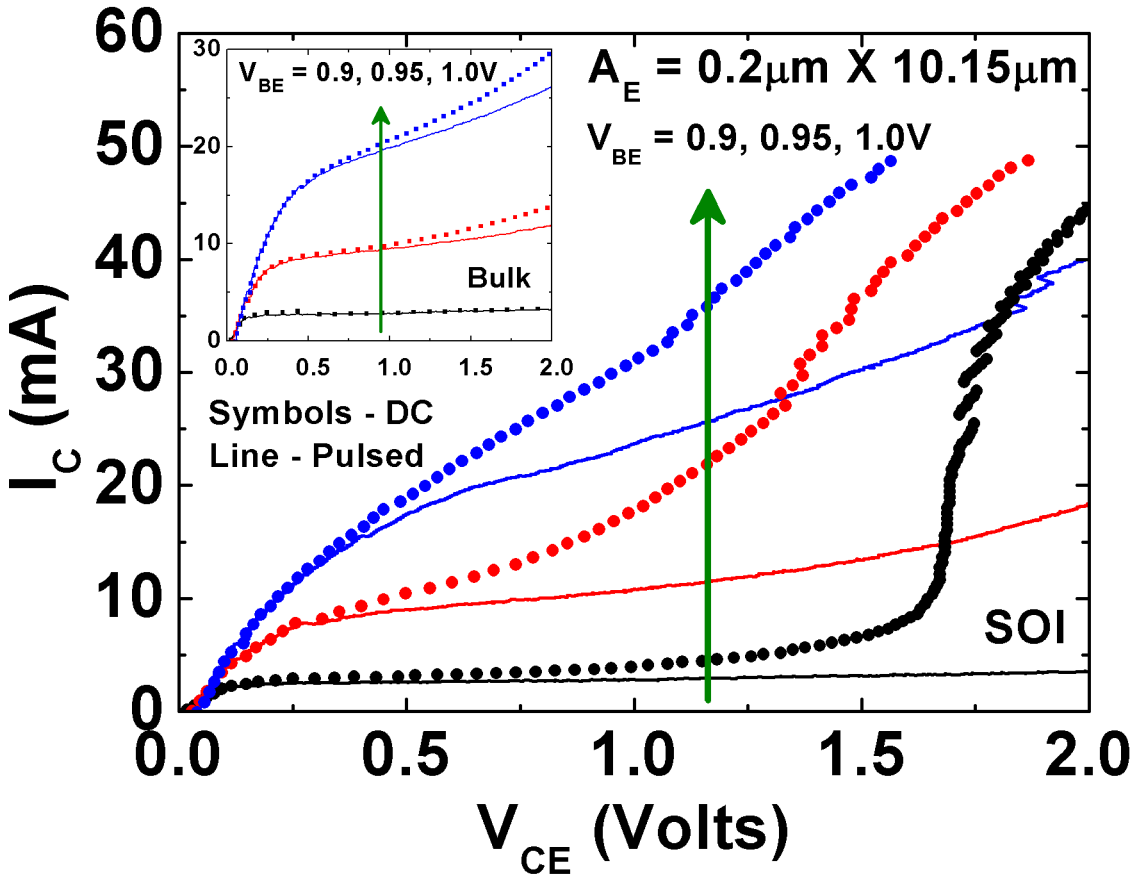

Figure 21: Forced-V_{BE} output characterisitics from DC and pulsed measurements for the Bulk (inset) and SOI device [3].

In order to further understand the implications of these observed SH effects in SiGe HBTs on SOI and how it will evolve with further scaling, a 200 GHz npn SiGe HBT (with STI and DT) TCAD deck with profiles optimized for bulk-Si substrate was simulated on a thick-film SOI substrate. In Fig. 22, Gummel characteristics from full ET simulations under FVB condition (using Sentaurus Work Bench from Synopsys, Inc. [56]) show TR in I_C and I_B at a critical V_{BE}, which progressively moves to lower V_{BE} with increasing V_{CB} or SH. This results in a sharp collapse of β_{DC} at the onset of TR. No similar SH effects were

observed in the TCAD simulations of the bulk devices. This shows that SH effects related to positive ET feedback will continue to impose maximum allowed voltage limitations on the safe operation of scaled SiGe HBTs on SOI.

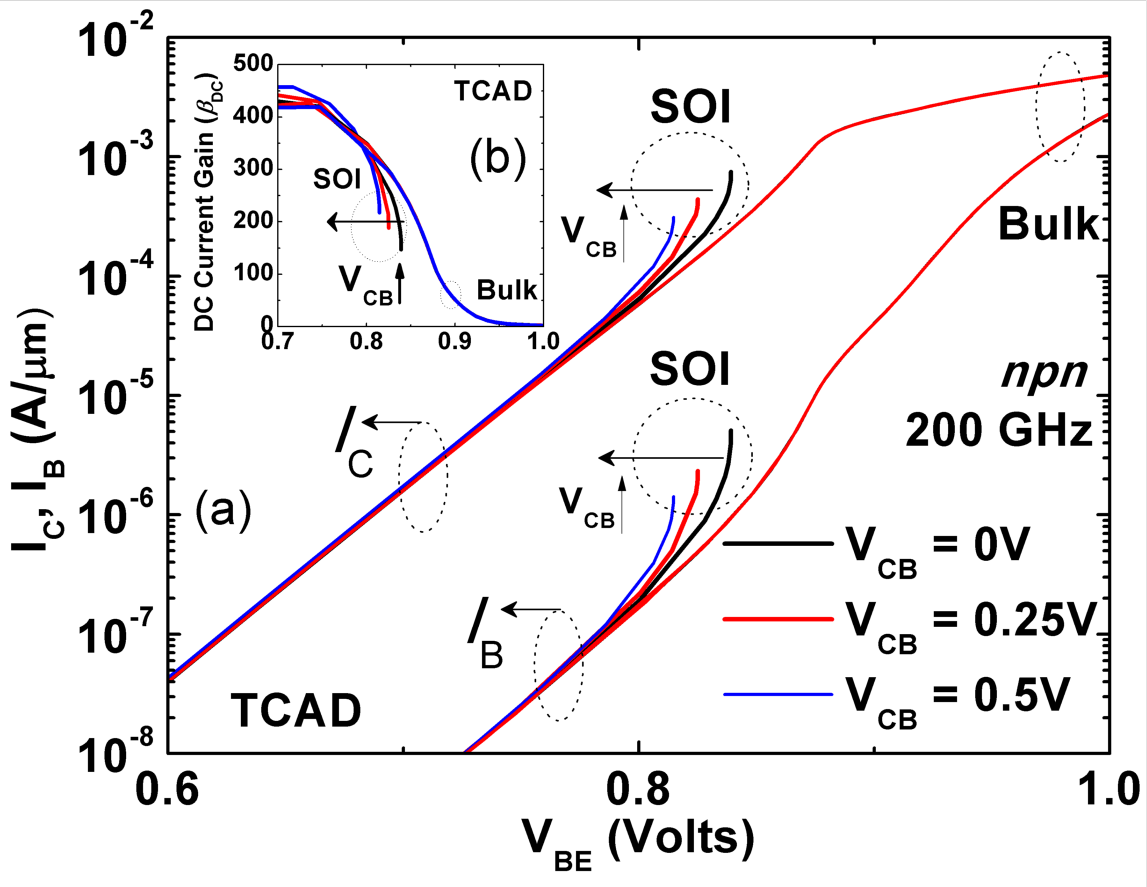

Figure 22: Evolution of the DC (a) Gummel Characteristics and (b) Current Gain plots with increasing V_{CB} (from full electro-thermal TCAD simulations for a 200 GHz npn SiGe HBT) [3].

4.3 A Hot-Carrier Reliability Issue for Thick-Film SOI SiGe HBTs

As a part of the DC characterization, when gummel characteristics were measured for increasing V_{CB} values on both the SOI and bulk device, a key reliability issue was observed as shown in Fig. 23. As self-heating increase within the devices at higher V_{CB}, the device temperature for the SOI device will be much higher compared to the bulk device. The V_{CB} values and V_{BE} values were limited to 1.0 V to reduce impact-ionization. It was observed that even with a weaker impact ionization coefficient (due to higher device temperature), the

SOI device exhibits increasing leakage at low-injection, a signature of hot-carrier damage at the BE junction spacer oxide-silicon interface. At the same time, the bulk device do not exhibit any observable change in the leakage at low-injection on the gummel plot. Since impact-ionization can be ruled out as the source of hot-carriers causing the damage in the SOI device (since impact-ionization is much weaker compared to the bulk device which does not see any damage), other sources of hot-carriers needs to be counted.

It is known from literature that both impact-ionization and auger-recombination can generate hot-carriers capable of causing damage in bipolar transistors [74]. The current in these devices will be constricted towards the center due to thermal instabilities, leading to a very high current density and consequently a strong auger recombination process. As opposed to the impact-ionization process which has a negative correlation with temperature, auger-recombination is strongly enhanced with temperature. Hence the auger-recombination process will be much stronger in a device with higher temperature or self-heating, thereby indicating that hot-carrier density generated by auger-recombination will also increase accordingly with self-heating or higher device temperature. The damage created within the SOI device can thus be attributed to hot-carriers generated by Auger-recombination, in the absence of any observable impact-ionization damage.

This mechanism will have strong impact on the hot-carrier reliability of SOI devices which will be enhanced at higher temperatures, or in any aggresively scaled SiGe HBT device with significant self-heating. In a nutshell, this bears significant implications for high-temperature operation or high-current operation of the SOI devices, thereby reducing the hot-carrier induced SOA of the device.

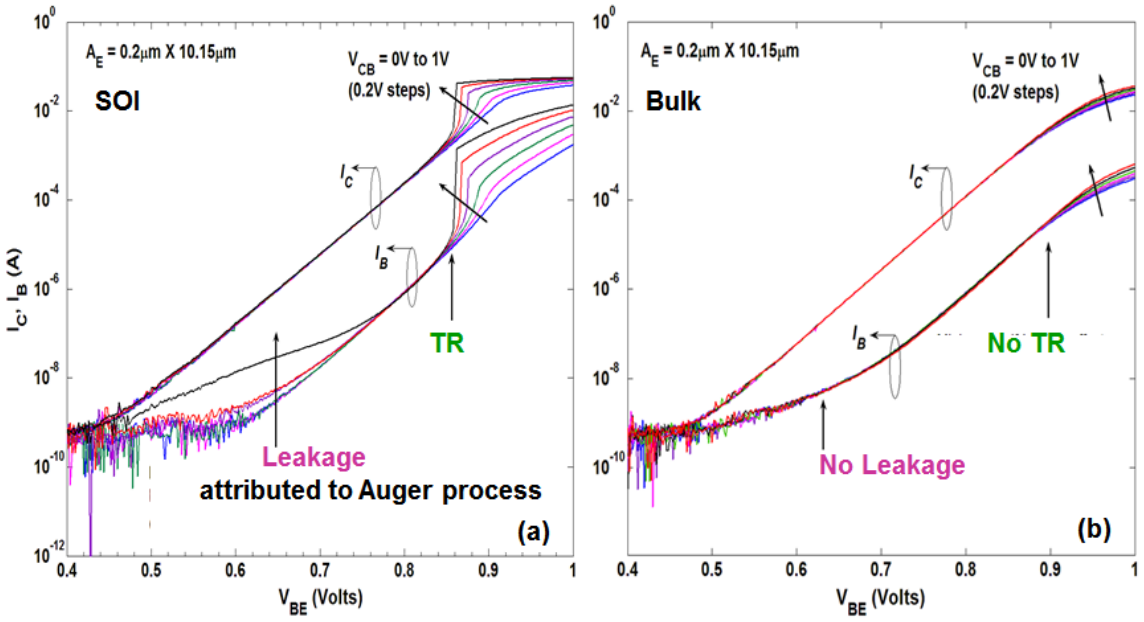

Figure 23: Evolution of the DC Gummel Characteristics for (a) the SOI and (b) the bulk SiGe HBTs when V_{CB} values are swept to increase self-heating. The SOI device shows an increasingly higher leakage from hot-carrier degradation resulting from the sweep, when compared to the bulk device which shows no excess leakage

4.4 Collector Doping Dependence of Thermal Resistance in SiGe HBTs

The device structures on thick-film SOI available for this study also included SiGe HBTs with different breakdown voltages, resulting from variations only in their selectively-implanted-collector (SIC) doping. Everything else regarding the doping and Ge profiles, as well as the device physical structure was similar. The previous section presented results on the high-speed device with the lowest breakdown voltage, which also has the highest SIC doping. However, it was observed from measurements that under similar V_{CB} (or V_{CE}) conditions, the devices with the highest breakdown voltage (or lowest SIC doping) shows TR similar to the high-performance devices as depicted in Fig. 21. A higher-breakdown device is expected to have lower self-heating at similar operating voltage and current values, due to a lower collector-base junction electric field. However, in this case, the device showed thermal runaway for all device sizes (varying in emitter lengths). Thermal resistance (R_{TH})

measurements [29, 30] on the high-breakdown devices gives a higher R_{TH} value over high-performance devices, while keeping the device emitter geometry same. This trend was validated across three different device sizes on SOI, and even for 2 different device types with a variable SIC doping from a higher performance technology on bulk silicon. Results confirmed that the high-performance devices have a lower thermal resistance compared to the high-breakdown devices for the same emitter geometry and process technology (as shown in Tables 3 and 4 below). Since most of the heat dissipation in SiGe HBTs occurs through the bottom of the wafer, the difference in R_{TH} becomes more significant for smaller device sizes. This difference in thermal conductance can be attributed to the collector-doping concentration based on reported literature, which typically mentions about doping dependence of the thermal conductivity in silicon [75]. The reported trends are fairly skewed in literature for epitaxial silicon films. However, this is the first report of a lower R_{TH} from higher SIC doping in SiGe HBTs. A higher carrier concentration in the SIC region leads to more interactions with phonons, which are responsible for the heat transfer across a material. Thereby, self-heating is a consequence of the difference between heat-dissipation and heat-conduction, which leads to TR for both types of devices. Further insights can be derived from TCAD simulations with phonon dispersion relationships on similar device structures. Overall, this has significant implications for device design and performance scaling, where increased collector doping will be expected to help reduce the thermal resistance of the aggresively scaled higher-performance devices both in SOI and Bulk technologies.

Table 3: R_{TH} dependence on SIC doping and geometry for the SOI technology

Device Size (A_E)	**R_{TH} (High-Performance)**	**R_{TH} (High-Breakdown)**
0.2µm × 1.7µm	11.826 × 10^3 K/W	15.953 × 10^3 K/W
0.2µm × 4.52µm	7.739 × 10^3 K/W	9.6475 × 10^3 K/W
0.2µm × 10.15µm	5.1971 × 10^3 K/W	6.067 × 10^3 K/W

Table 4: R_{TH} dependence on SIC doping and geometry for the bulk technology

Device Size (A_E)	R_{TH} (High-Performance)	R_{TH} (High-Breakdown)
$0.1\mu m \times 10\mu m$	2.7209×10^3 K/W	4.7547×10^3 K/W

4.5 Effects of Electrothermal Constraints on the RF Linearity of Scaled Thick-Film SOI devices

Since stronger self-heating in the TF-SOI devices leads to a higher junction temperature compared to the bulk devices, it leads to stronger non-uniformities in the current distribution across the emitter junction area and consequently a different potential distribution across the BE junction. Both of these factors causes the TF-SOI device to have different device characteristics compared to the bulk (as studied in the previous sections) and hence it is important to explore how consequently the RF linearity of these devices vary under both minimal and strong self-heating conditions. The linearity response of the two SiGe HBT devices (TF-SOI and bulk) are presented and compared here. This study is motivated by the fact that self-heating can potentially couple with the linearity performance FoM for the transistors through differences in the current and potential distributions across the two devices, which has similar doping and 2-D cross-sections to begin with.

The most commonly used device figures-of-merit for RF linearity are the 1-dB compression point, P1dB (a measure of large-signal linearity), and the third-order intercept point, TOI (a measure of small-signal linearity). TOI is expressed as either the input-referred (IIP3) or output-referred (OIP3) third-order intercept point. P1dB indicates the input power level (P_{IN}) that causes the small-signal gain (i.e., S_{21}) at a specific RF frequency to drop by 1-dB from its maximum value at lower P_{IN}. OIP3 (or IIP3) is the extrapolated output (or input) power where the fundamental tone and the third-order intermodulation product power has the same magnitude at a specified RF frequency. The presented work

investigates both large-signal and small-signal linearity of the TF-SOI and bulk transistors. Analysis of the differences in the above mentioned linearity FoM will provide insight into how strong self-heating effects can influence the linearity of a TF-SOI transistor, and consequently changes RF SOA of the device.

The S-parameters measurement setup (displayed in Fig. 24) consists of a network analyzer (Agilent E8363B PNA), bias-tees, input and output tuners, Keithley 2400 source-measurement meters, and the on-wafer RF prober. All measurements were performed at room temperature (300 K). For both single-tone (P1dB) and two-tone linearity (OIP3) measurements, the experimental setup included the following equipment: two signal generators (Agilent HP83732A), two *dc* supplies (Keithley 2400 SMU), three isolators (Ditom model D3108), one combiner (Narda model 30183), two bias-tees (Anritsu model K251), one power meter (Agilent E4419), two bi-directional couplers (Krytar model 2618), two Focus Microwaves programmable tuners (model CCMT-1818), and one spectrum analyzer (Agilent E4407B). The instruments and tuners were controlled using the Focus Microwaves Load-Pull Explorer software application program.

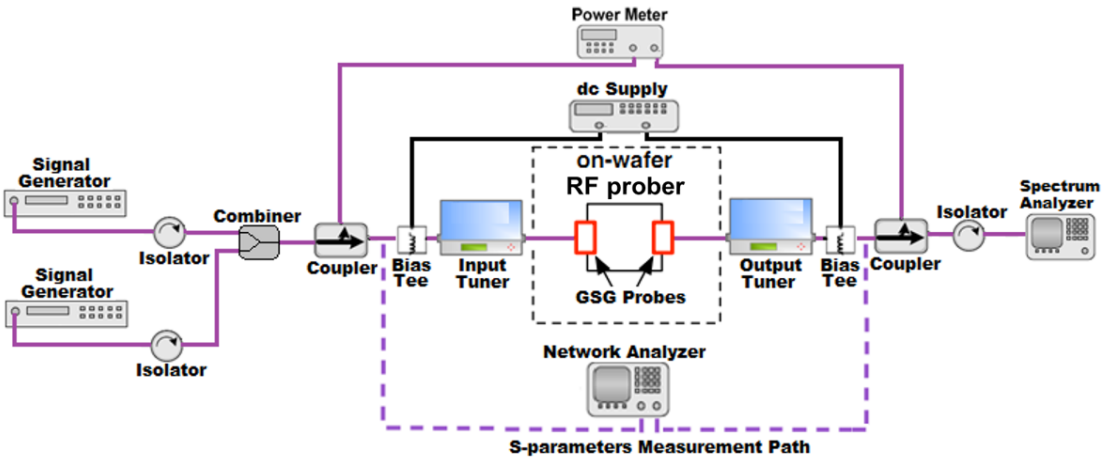

Figure 24: The experimental setup for on-wafer linearity measurements

To ensure accurate power readings for the linearity measurements, calibration loss files were generated for the input and output paths using power flatness and multiple reference planes calibration techniques. Two different sets of measurements were performed, first

at 50 ohm impedance and the other with impedance matching at the input and output for maximizing output power and minimizing loss. Before running the second set of linearity measurements with impedance matching, source- and load-pull were performed for both the TF-SOI and bulk devices at 8 GHz and 18 GHz using a constant P_{IN} set to -20 dBm. The source and load impedances were tuned for maximum output power (P_{OUT}) at each frequency and bias. The results presented here were fairly representative of multiple devices. High-speed (HS) devices with an emitter area A_E of 0.2×10.15 μm^2 were used in this study for both the SOI and bulk variants. All *ac* device test structures were characterized in a common-emitter (CE) mode. The devices were biased in the forced-I_B mode for the 50 ohm measurements to reduce instabilities from thermal runaway. However, for the linearity measurements with matched impedance, devices were biased in the forced-V_{BE} mode.

4.5.1 Analysis of the Linearity Measurements with 50 ohm impedance

The large-signal linearity (P1dB) measured for the SOI and bulk devices for different relevant I_C values under three V_{CE} values (of 1.0, 1.5, and 2.0 V for varying self-heating conditions) at 18 GHz is shown in Fig. 25. As shown for both devices, there is clearly very less variation in P1dB values at both frequencies and three V_{CE} values except at the highest I_C values close to peak f_T uded for the measurements. Under minimal self-heating conditions (at V_{CE} = 1.0 V), P1dB values shows a maximum variation of ~5 dBm at closest to the peak f_T values. The results do not indicate any degradation in the large-signal linearity (P1dB values) due to self-heating in the SOI devices under forced-I_B operation. The small-signal linearity (OIP3) for the bulk and SOI devices at 18 GHz for the three different V_{CE} values are shown in Fig. 26. At I_C < 1 mA, the bulk device shows higher OIP3 compared to the SOI device. The differences then increase once the devices are biased close to the strong self-heating regions or thermal runaway (I_C > 3 mA). The trends for the bulk device is different for the highest I_C point where the measured OIP3 at V_{CE} = 1.5 V is the highest, while OIP3 is the highest at V_{CE} = 2.0 V for the SOI device. For the SOI device, OIP3 reaches a peak value before either decreasing or staying flat at at higher I_C (different from

the bulk device). The differences can be attributed to the g_m and C_{CB} non-linearities arising out of the non-uniformities in the current distributions, further leading to self-heating related bias instabilities. At V_{CE} = 2.0 V, even when TR is observed, the measured OIP3 for the SOI device is the highest, confirming that there is no significant degradation in the RF SOA of the devices due to strong self-heating, when measurements are done with a 50 ohm system impedance.

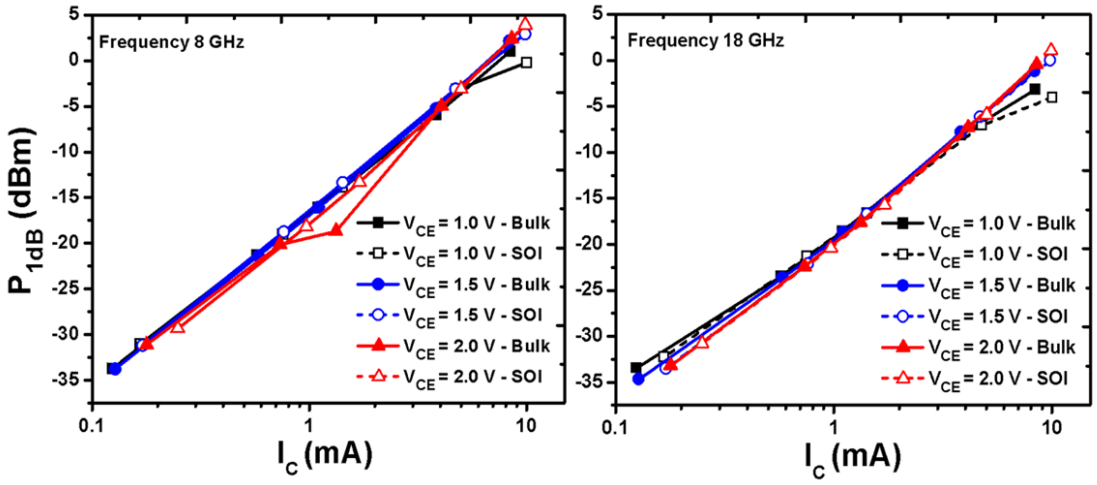

Figure 25: Large-signal linearity FoM (P1dB) compared for the SOI and bulk devices under three V_{CE} values (1.0, 1.5, 2.0V) at two different RF frequencies, (a) 8 GHz and (b) 18 GHz.

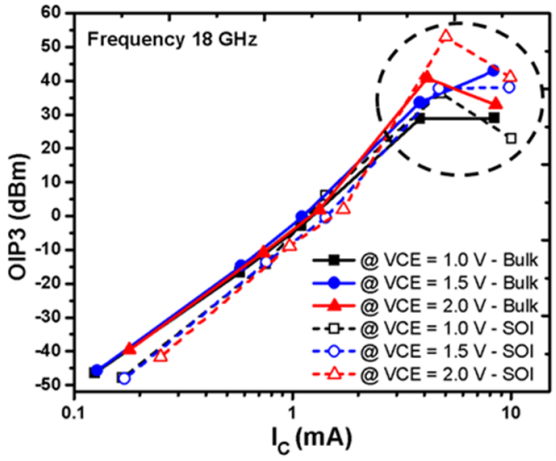

Figure 26: Small-signal linearity FoM (OIP3) compared for the SOI and bulk devices under three V_{CE} values (1.0, 1.5, 2.0V) at 18 GHz.

4.5.2 Analysis of the Linearity Results with Impedance Matching

The small-signal linearity at 8 GHz for different I_C values from medium to high injection at V_{CE} = 1.5 V (with strong self-heating, but minimal thermal runaway) is shown in Fig. 27. These results are representative of measurements on multiple die. It can be observed from the data that for all the J_C values considered, the SOI device has a lower OIP3 compared to the bulk device in this bias mode. The difference is fairly consistent at low to medium injection levels (with minimal self-heating). However, at close to and above peak f_T (J_C ~ 5mA/μm^2), where self-heating is dominant the difference in OIP3 between the two devices increase. However for the 16 GHz case, the SOI and bulk differences are minimal until strong self-heating sets in. Taken together, this implies that SOI devices will have a lower linearity even when self-heating is minimized, due to more dominant capacitive non-linearities resulting from oxide vs. junction isolation. However, with stronger self-heating, at close to or above peak f_T, that difference widens, which can be attributed to increase in both g_m and capacitive non-linearities resulting from the stronger temperature gradient across the device. This temperature gradient results in non-uniform current and potential distributions, leading to bias instabilities [Curtis Grens, Phd Dissertation, Georgia Tech].

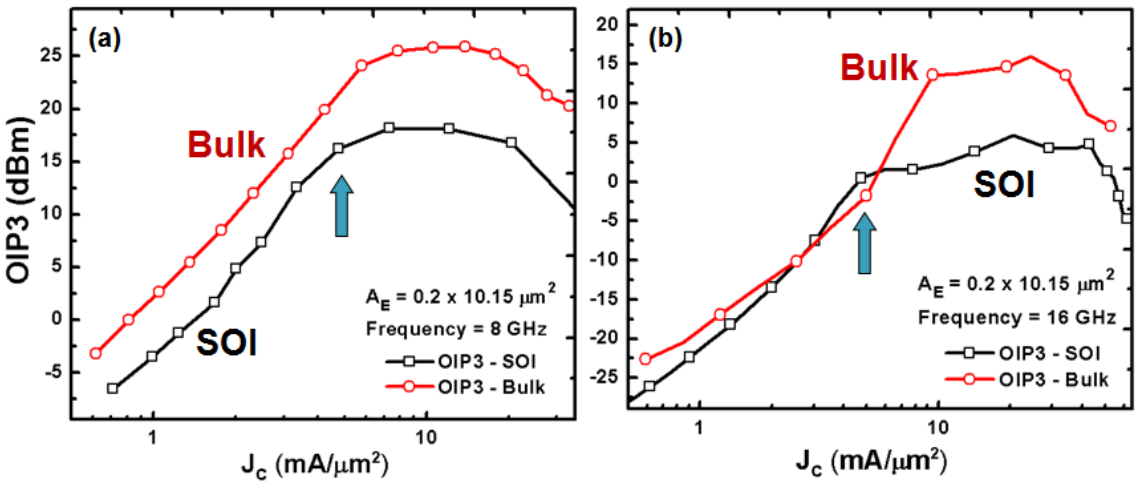

Figure 27: Small-signal linearity FoM (OIP3) compared for the SOI and bulk devices at (a) 8GHz and (b) 16 GHz under forced-V_{BE} biasing with V_{CE} = 1.5V.

This implies a significant reduction in the RF SOA of the TF-SOI devices with forced-V_{BE} bias in the CE operation mode, when compared to those fabricated on bulk with identical profiles and geometry. Although further results are needed to make more generalized conclusions, these results indicate that RF SOA of TF-SOI devices can be reduced based on the bias and topology of the circuit application where the devices are used.

4.6 Summary

It has been clearly demonstrated that strong SH effects resulting from positive ET feedback will continue to impose serious fundamental electrical limitations on device operation for highly-scaled SiGe HBTs on SOI. Different device optimization, processing and thermal management techniques will be required in order to continue scaling high-performance SiGe HBTs in an SOI BiCMOS process. Electrical and thermal isolation techniques need to be decoupled during technology development. As shown here, different circuit biasing techniques may be required to prevent TR. In other words, constraints on allowed bias current or voltages for reliable operation will be determined by thermal rather than electrical limitations. This will entail incorporating new figures-of-merit for ET constraints in device compact models for circuit simulation. The only factor that might mitigate self-heating effects in scaling SiGe HBTs on TF-SOI could come from the fact that R_{TH} of these devices decrease with increased collector doping. This important piece of information can be used for designing SiGe HBTs with very high performance both on bulk and SOI.

The RF linearity measurements and comparisons between SOI and the Bulk devices clearly indicate that SOI devices could potentially have lower linearity based on how they are biased in a circuit application, leading to a reduction in the RF SOA of the devices. Taken together, this implies the need for technological advancements to mitigate these effects, in addition to efforts for modeling the linearity differences more predictively.

CHAPTER 5

PREDICTIVE OUTPUT CONDUCTANCE MODELING OF SIGE HBTS

The presented investigation provides insight into the issues involved in accurate TCAD modeling of V_A in advanced 200 GHz (3rd generation; 3G) SiGe HBTs. Experimental results are also presented to verify our assertions. For the first time, a scaling-induced divergence between V_A predictions from TCAD and data, due to self-heating (SH) effects [4].

5.1 Output Conductance Modeling Issues Observed from TCAD Simulations

2-D TCAD simulations were performed to design and optimize a new 3G *npn* SiGe HBT using the Sentaurus Workbench (SWB) commercial simulation environment [1, 56]. The optimized *npn* SiGe HBT at the 130 nm lithography node was found to have a peak f_T and $f_{MAX} \geq 200$ GHz, $BV_{CEO} \sim 1.8$ V. All of these FoM are comparable to the values reported in the literature for 3G *npn* SiGe HBTs. However, the V_A extracted for the same optimized device deck was found to be surprisingly low (< 20 V), for both forced-I_B and forced-V_{BE} input drive simulations at low-to-medium injection levels, ruling out any role of inaccurate neutral base recombination modeling in the extracted values. To probe this further, an available calibrated TCAD deck for an existing commercial 200 GHz *npn* SiGe HBT [76] was also examined, and this TCAD deck also showed an unrealistically low V_A compared to measured values (< 20 V simulated vs. 110 V measured). However, for a similar calibrated deck for a 50 GHz (1st generation; 1G) commercial *npn* SiGe HBT [77], the simulated V_A value (100 V) closely matched the measured value (90 V). To explore the origin of this apparent scaling-induced difference in simulated vs. measured V_A, we modified our 2-D TCAD deck to include self-consistent SH equations. Careful consideration was given to the thermal contact placement, thermal resistance, boundary conditions, and optimization

for robust convergence. The V_A (forced-I_B) extracted from the electro-thermal simulation was much higher compared to the isothermal simulation, clearly indicating the increasing importance of SH in a 3G SiGe HBT. As expected, the effect of SH on I_C becomes more significant at higher injection levels (Fig. 28). Realistic values of the thermal resistance must be used to achieve the correct I_C-V_{CE} slope (Fig. 29).

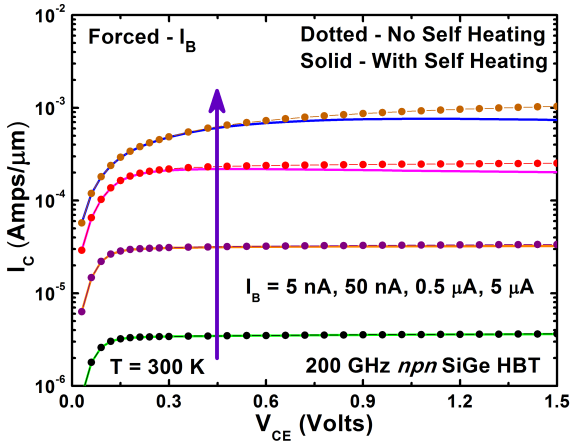

Figure 28: Simulated V_A extracted at I_B = 50 nA increase from 19 V to > 50 V after including SH [4].

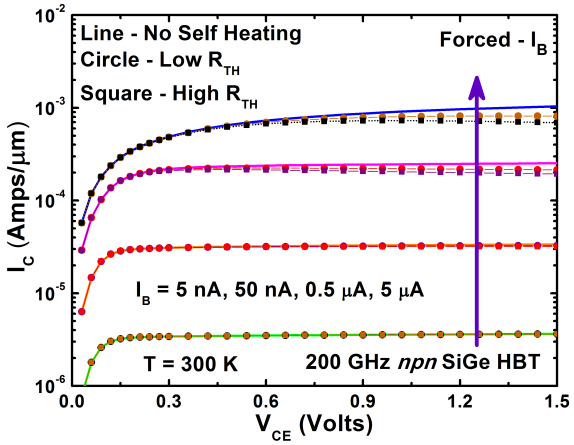

Figure 29: Simulated output characteristics of the optimized *npn* SiGe HBT with different thermal resistance at the top of the wafer [4].

5.2 Experimental Analysis of TCAD Results

To experimentally analyze the effects of SH in *npn* SiGe HBTs [3], the device output characteristics were measured using both *dc* and pulsed $I-V$ test systems. Pulse width (T_W)

and pulse separation (T_S) times were both varied to control SH during the measurements. As seen in Fig. 30, for shorter T_W, the SH signature is weak, but becomes more apparent for longer T_W, as seen from the negative slope in the I_C-V_{CE} plot. In Fig. 31 it is evident that even for the same T_W, a long enough T_S allows the device to "cool-down" between pulses, indicating that the thermal time constant of the device is less than about 10 ms.

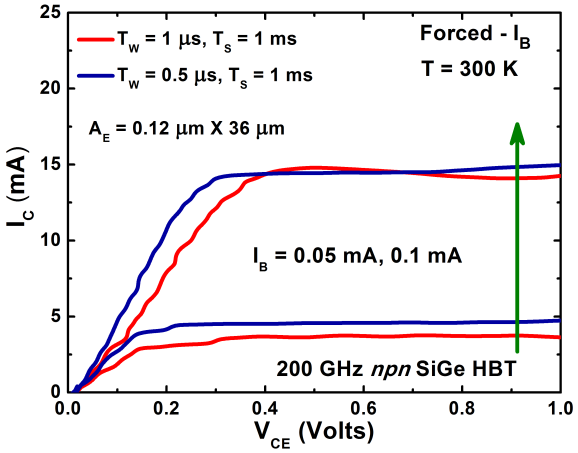

Figure 30: Measured forced-I_B output characteristics with varying T_W at constant T_S [4].

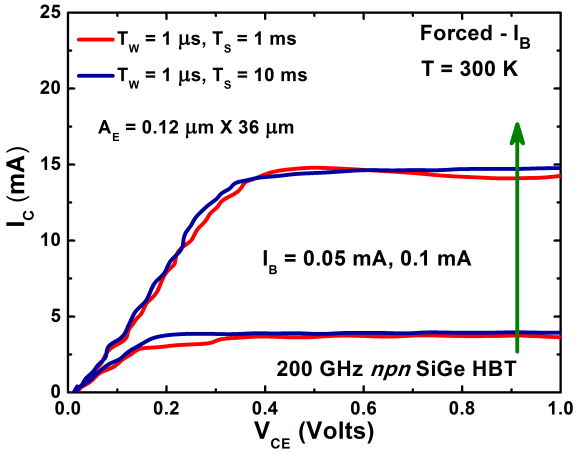

Figure 31: Measured forced-I_B output characteristics with varying T_S at constant T_W [4].

Fig. 32 shows that the V_A (forced-I_B) measured using T_W = 1 s, T_S = 10 ms is much lower than the V_A measured under *dc* conditions, consistent with our TCAD predictions. Pulsed measurements on 1G SiGe HBTs, however, do not show the same sensitivity to SH at low-to-medium injection, as seen in Fig. 33, again consistent with the TCAD results.

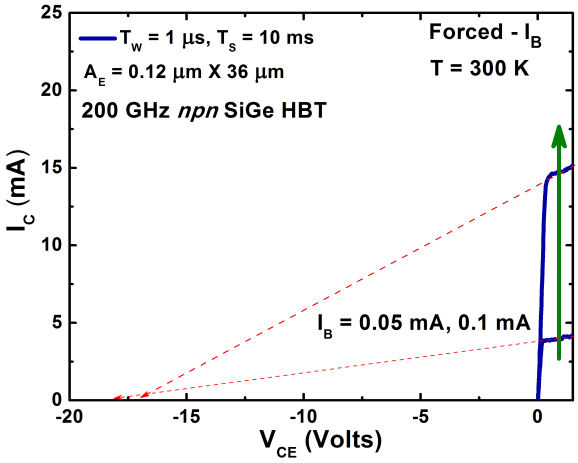

Figure 32: V_{A} extracted from forced-I_{B} pulsed measurements (to minimize self-heating) is lower compared to dc measurements [4].

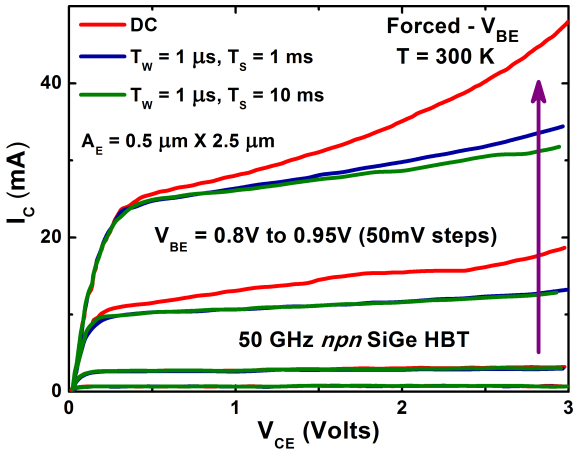

Figure 33: Measured forced-V_{BE} output characteristics compared between *dc* and pulsed measurements [4].

5.3 Summary

Taken together, these TCAD modeling results and analysis and the experimental verifications, suggest that there are significant implications for using TCAD for predictive modeling of output conductance in SiGe HBTs at aggressive scaling nodes. As the devices gets scaled to higher performance nodes, self-heating within the device will become increasingly more dominant based on higher current densities and electric fields. In essence, it will become more important to include self-heating as device performance gets scaled.

CHAPTER 6

PREDICTIVE TCAD MODELING OF THE SCALING-INDUCED, REVERSE-BIASED, EMITTER-BASE TUNNELING CURRENT IN SIGE HBTS

The presented study here attempts to fill the gap mentioned earlier by investigating the nature of the reverse-biased EBJ tunneling current and its mechanisms in three distinct generations of *npn* SiGe HBTs (50, 100, and 200 GHz) and in two generations of *pnp* SiGe HBTs (100 and 200 GHz), using a combination of both measurements and TCAD modeling, and by further examining the relation between RB EBJ stress in the tunneling-dominated regime to device reliability as a function of performance scaling [5].

6.1 Experimental and Simulation Details

The physical cross-section of the bulk-Si SiGe HBT profile used for both the measurements and TCAD simulations has an EB spacer, STI, DT, SIC, etc. Three generations (50, 100 and 200 GHz) of *npn* SiGe HBTs and a 2nd generation (100 GHz) *pnp* SiGe HBT were used for the dc measurements [19,76–78]. Optimized candidate device profiles for three generations of *npn* SiGe HBTs (50, 100, and 200 GHz) and two generations of *pnp* SiGe HBTs (100 and 200 GHz) were used for TCAD simulations to enable analysis of the correlation to the measurement results [1]. Finally, accelerated stressing of the RB EBJ in the tunneling regime of the *npn* devices was performed to examine its relation to the nature of the device degradation and reliability. For the purposes of this study, to optimize the simulation time due to challenging convergence issues, RB EBJ tunneling simulations have been typically limited to an EBJ V_R of 1.0 V. RB EBJ current measurements were mostly limited to a V_R of 1.0 V to prevent EBJ degradation, which will change the tunneling current itself.

DC measurements were performed using an Agilent 4156B Semiconductor Parameter Analyzer. All measurements were performed at room temperature. The typical device bias configurations used for the RB EBJ measurements and simulations are shown in Fig.

34. The EBJ was RB using an emitter bias while the collector and base terminals were grounded for both the *npn* and *pnp* devices. Emitter geometries with minimum width (W_E) and different lengths (L_E) were examined based on their availability in the different process technologies. Device simulations were performed using the Sentaurus Workbench (SWB) tool from Synopsys [56]. Models used for the EBJ tunneling simulations include a non-local BTBT model (developed in [79]) with a non-local mesh, and the Schenk TAT model [41]. Standard models for Shockley-Read-Hall (SRH) and Auger recombination were used. Unless otherwise noted, DD carrier transport was used in the rest of the device, while electron and hole tunneling masses (m_t) of $m_c = 0.2m_0$ and $m_v = m_0$ respectively, were used, as reported in the literature [80].

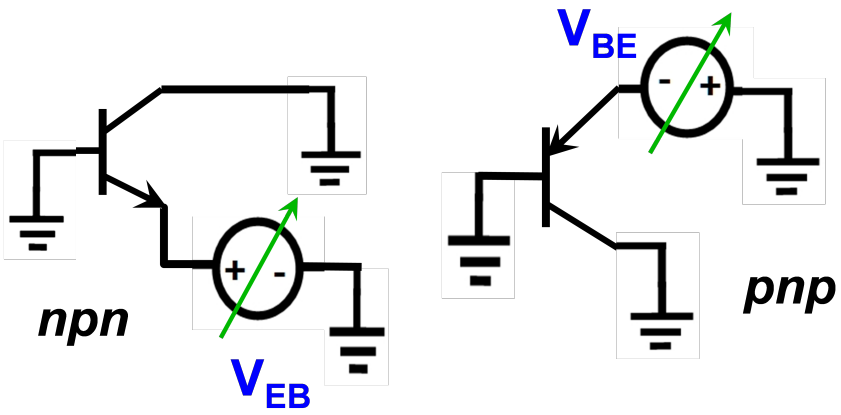

Figure 34: Device bias configurations used for the tunneling measurements and simulations on the RB EBJ of the SiGe HBTs [5].

6.2 Results from a 200 GHz *npn* SiGe HBT

Results from the RB EBJ device simulations performed on a 200 GHz *npn* SiGe HBT profile with and without the BTBT model are shown in Fig. 35(a). $E_{0,\text{peak}}$ for the device profile is 1.16106 V/cm, high enough for both BTBT and TAT mechanisms to be operative [41]. This result helps in estimating the BTBT contribution to I_B from the RB EBJ in the device, which for the measured device is found to be significant. Simulation results shown in Fig. 35(a) indicate that the RB EBJ current for the measured 200 GHz *npn* is well-accounted for by the BTBT mechanism, as expected from the theoretical models [40, 41]. The validity of

using a DD transport model was examined by comparing it with simulation results from a HD transport model for the same device, as shown in Fig. 35(a). The BTBT contribution to I_B from both transport models matches well for the bias levels considered in this study. However, the HD model shows much more severe convergence issues at a lower V_R as well as a significant increase in the simulation time, thereby justifying the use of DD transport for the rest of this study. Based on their influence on the BTBT and TAT mechanisms, important factors considered for these simulations include effects from varying material parameters (e.g., bandgap) for the SiGe region of the device (in the absence of full process simulation flow to include the stress effects), and the effective tunneling masses (m_c, m_v). When a Monte-Carlo generated parameter file for the SiGe region (available within SWB) is used (as opposed to the normal parameter set that interpolates between Si and Ge properties), the effect on the BTBT component of I_B from an effective change in the bandgap is shown in Fig. 35(b). The impact of varying tunneling masses (m_c, m_v) on the BTBT current is also shown in Fig. 35(b). Although the BTBT contribution to I_B is not significantly affected by a change in the parameter file, it is more sensitive to changes in the tunneling masses. For all simulations presented in the remainder of this study, a normal parameter file was used, unless otherwise noted.

It was observed that the measured I_B from the RB EBJ of the 200 GHz *npn* scales linearly with the emitter area of the device, implying that the areal component of the tunneling current is dominant and contributions from the emitter periphery are negligible [81]. This was also found to be the same for other devices considered in this study. The measured I_B is compared to RB EBJ simulations with only BTBT and with both BTBT and TAT in Fig. 36(a). The results clearly demarcate two regions: one at very low V_R (0.1 V) where TAT dominates, and the other region above 0.1 V where BTBT dominates. This is consistent with previous studies which showed that TAT dominates at lower EBJ electric fields, whereas BTBT dominates at higher electric fields [40, 41]. The simulated I_B for the 200 GHz *npn* HBT matches very closely with the measured data in the TAT regime,

and although it is higher in the BTBT regime, it follows the measured trend closely. Since the measured and simulated device profiles are not identical, but are only comparable in performance, the simulations and measurements compare reasonably well without profile calibration to data. For the simulated device, Fig. 36(a) also shows that the BTBT regime is more sensitive to differences in tunneling masses as compared to the TAT regime. The temperature dependence of the simulated tunneling current in Fig 36(b) highlights the different inherent temperature dependencies of the two tunneling mechanisms. TAT shows a somewhat stronger temperature dependence compared to BTBT [40]. These results form the foundation for the next part of this study.

6.3 Results from 50, 100 and 200 GHz *npn* devices compared

Results from I_B measurements on the RB EBJ for three generations of SiGe HBTs are shown in Fig. 37. Also in Fig. 37, the measured data are compared to the simulation results with BTBT (for profiles with comparable performance). It is clear from Fig. 37 that the BTBT component of I_B scales strongly with device performance. Although it is evident that BTBT is the most dominant tunneling mechanism in the 100 and 200 GHz devices and accounts for the measured data significantly, this mechanism fails to estimate the I_B or its trend for the 50 GHz device. This becomes clearer if we consider the $E_{0,peak}$ for the simulated device profiles (which are 1.16×10^6, 9.32×10^5 and 4.47×10^5 V/cm for the 200, 100 and 50 GHz devices, respectively) and compare them with values reported in the literature [40, 41]. $E_{0,peak}$ for the simulated 50 GHz device is less than what is required for BTBT to be operative at a V_R of less than 1.0 V.

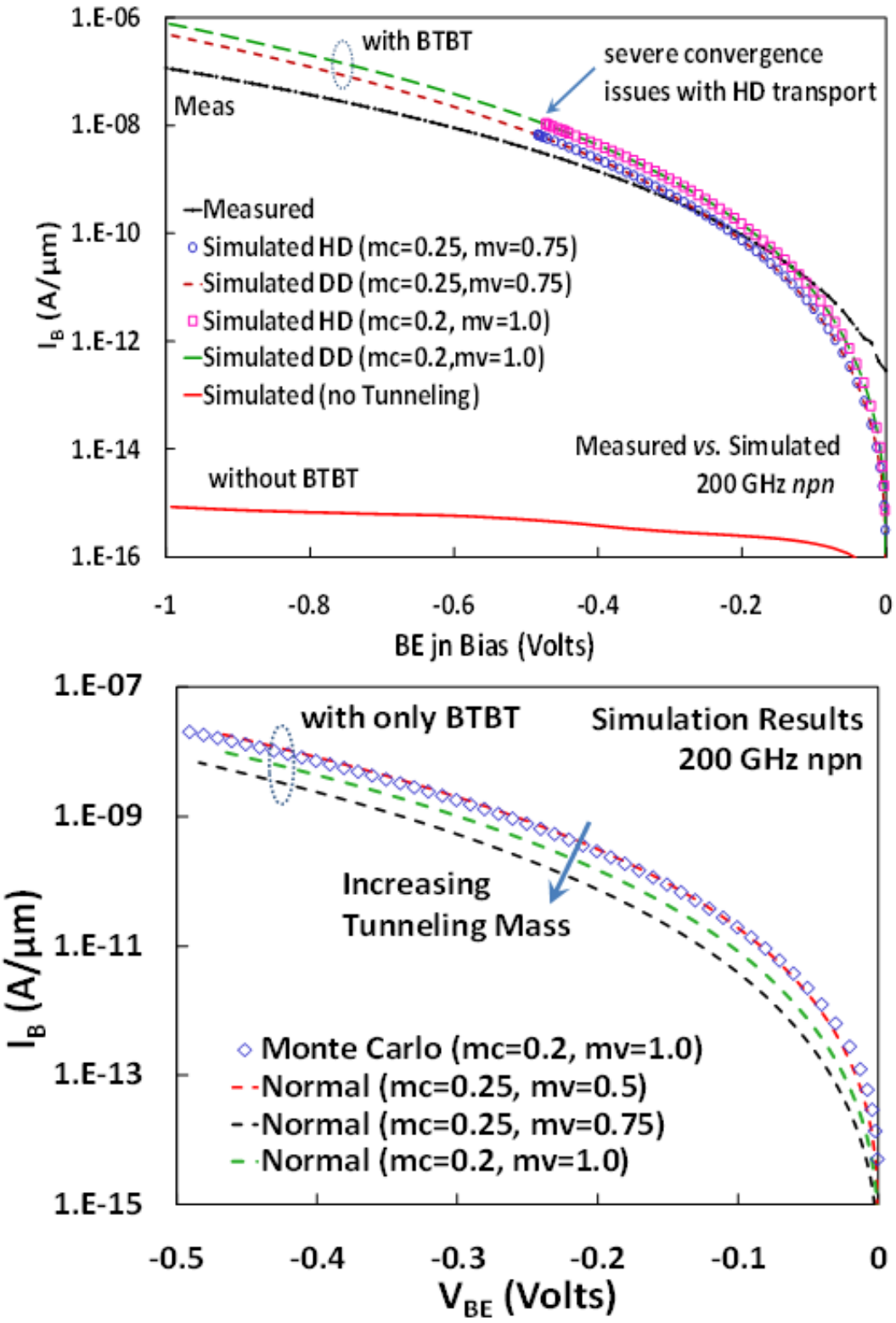

Figure 35: (Top, (a)) RB EBJ I_B from simulations with and without BTBT. Difference between HD and DD simulations for RB EBJ of a 200 GHz *npn* SiGe HBT profile, compared with measurement results from an actual device. (Bottom, (b)) Effect on simulated BTBT current component of I_B between two parameter files used, and the effect of tunneling mass variation on the BTBT current component of I_B in a RB EBJ [5].

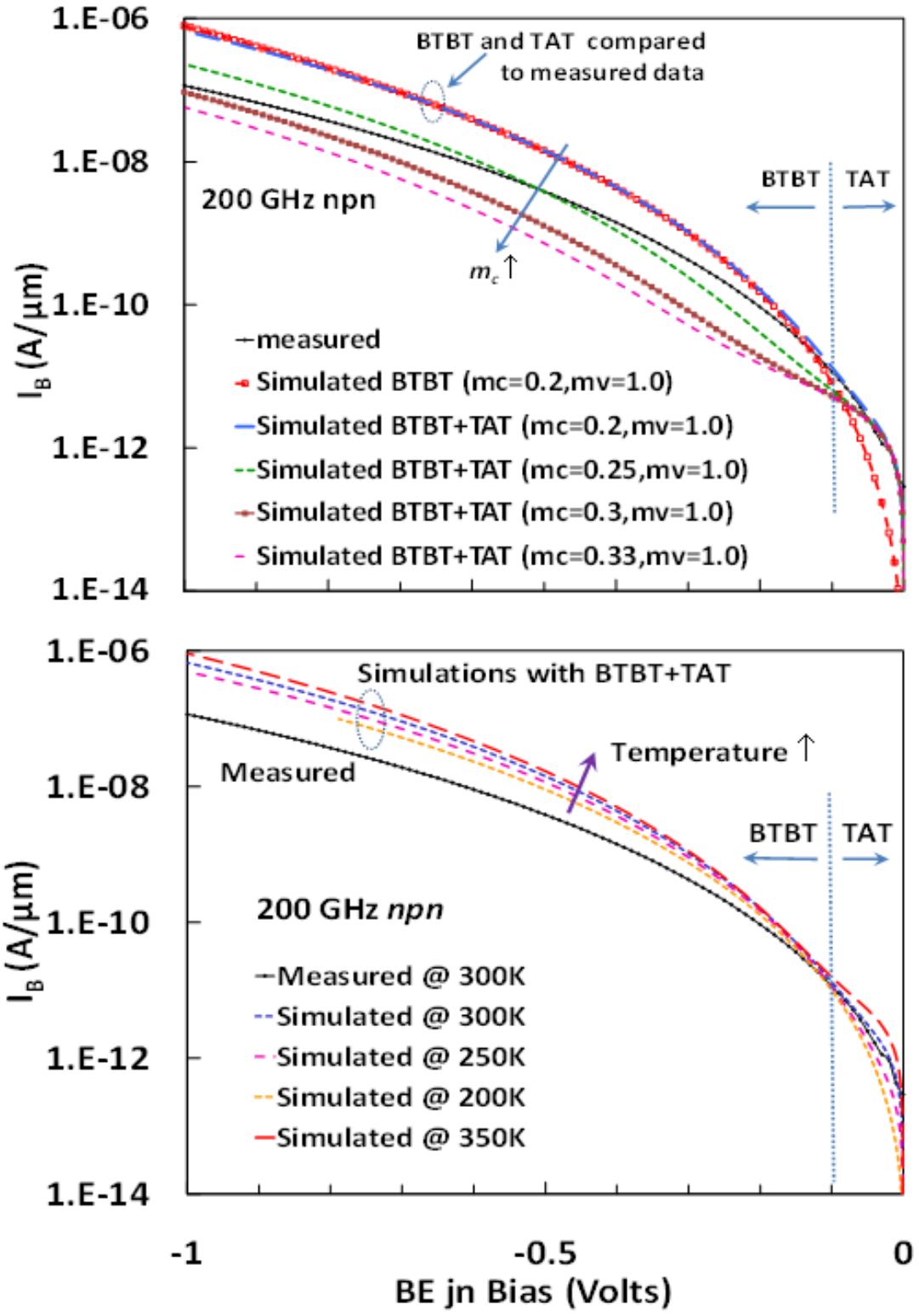

Figure 36: (Top, (a)) Simulated I_B from BTBT with and without TAT compared to measured data, and effect of tunneling mass variations on both the mechanisms. (Bottom, (b)) Simulation results for effect of temperature variation on both tunneling mechanisms [5].

When TAT is included with BTBT in the device simulator, the simulated I_B is compared with the measured data in Fig. 37. Although the measured and simulated I_B match well in the TAT dominated regime for the 200 GHz device, the simulated I_B overestimates the TAT current in the 100 GHz device when compared to the measured data. The simulated results for the 50 GHz profile show that the TAT component, when compared to only the BTBT current at a V_R of less than 1.0 V, accounts for most of the measured I_B, while underestimating it. Thus, Fig. 37 clearly shows that the BTBT model is more robust in predicting the current across different generations of SiGe HBT profiles, whereas a single set of parameters for the Schenk model would not be able to reliably predict the TAT current or its trend in the same devices. A previous investigation with similar tools that included simulation results on a SiGe heterostructure device has not commented on the increased measurement to simulation difference in the TAT dominated regime compared to the BTBT regime [80].

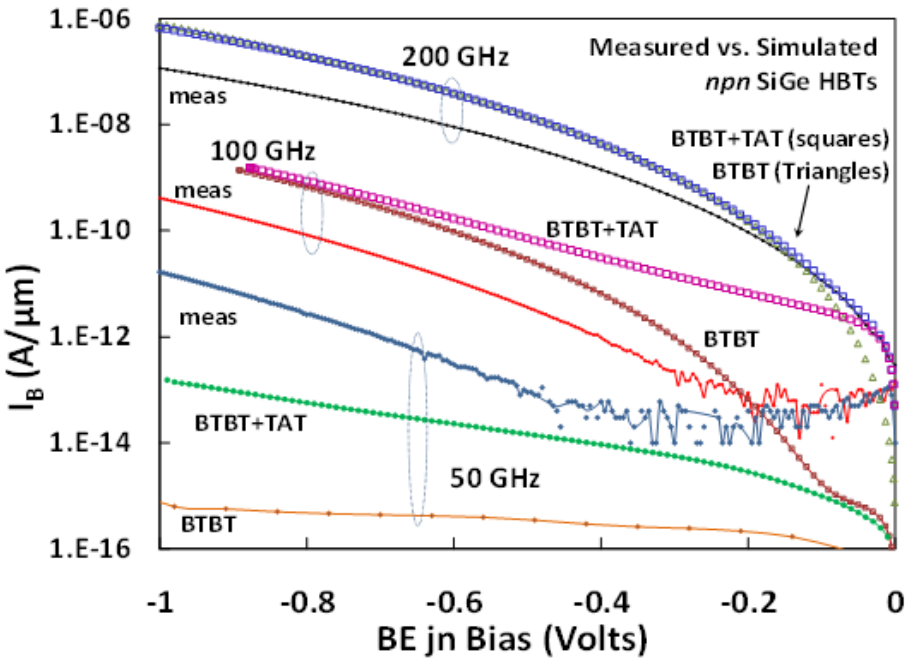

Figure 37: Simulation results with only BTBT, and with BTBT combined to TAT models are compared to the measured data for the 200, 100 and 50 GHz npn SiGe HBTs [5].

6.4 Results from 100 and 200 GHz *npn* and *pnp* devices compared

Results from the RB EBJ I_B measurements on a 100 GHz *pnp* device [19] are compared to simulations on an optimized 100 GHz *pnp* device profile in Fig. 38. Only the BTBT model was used for these simulations. Both measurements and simulations on 100 GHz *pnp* devices closely compare to the trend seen for the 100 GHz *npn* devices (for measured and simulated data). This further reinforces the robustness of the BTBT model for estimating the RB EBJ I_B in both *npn* and *pnp* devices, and shows that the model can also predict the tunneling current in complementary SiGe HBT platforms.

To take this a step further, an existing optimized 200 GHz *pnp* profile [1] was simulated with only BTBT current. Results are also shown in Fig. 38. The simulated results on the 200 GHz *pnp* compare very well with the 200 GHz *npn* devices, which is even more significant since they were optimized as 200 GHz complementary HBTs with matched performance.

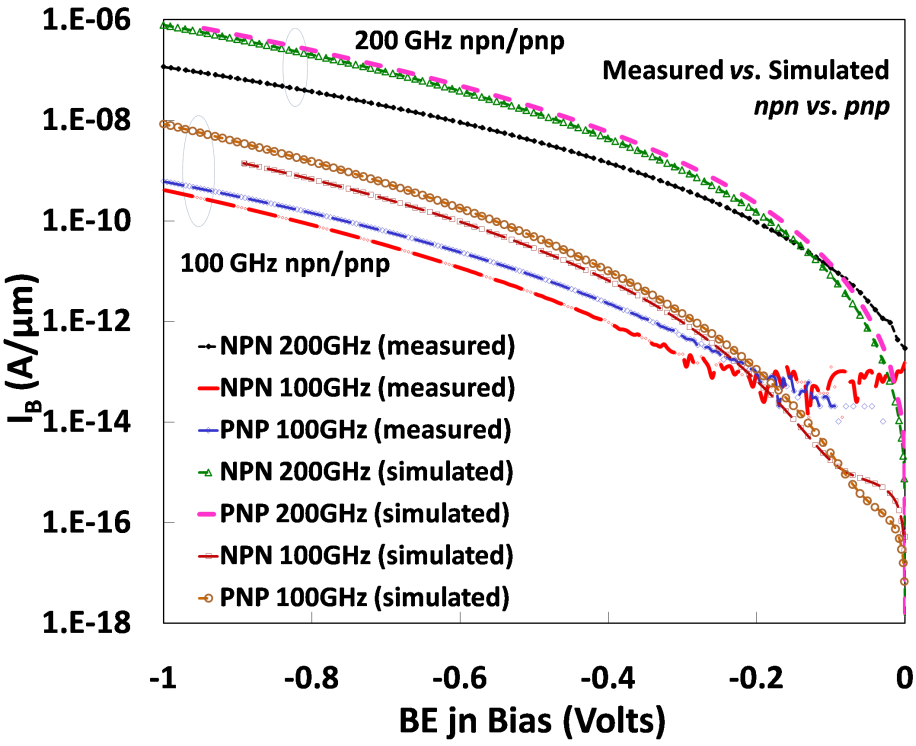

Figure 38: Comparison of the simulated RB EBJ I_B for optimized 100 GHz and 200 GHz npn and pnp profiles (with only BTBT). Measured I_B due to RB EBJ from 100 and 200 GHz npn and 100 GHz pnp devices are also shown for comparison to simulations [5].

6.5 Robustness to Tunneling Stress

Finally, accelerated stress in the tunneling regime of the EBJ with a constant reverse-bias (V_R) of 2.5 V and an open-collector (OC) configuration was performed for different generations of *npn* HBTs. Some of these devices showed a consequent increase in the non-ideal I_B component of their forward Gummel, as expected. The post-stress non-ideal I_B was then measured and compared to the pre-stress value at a pre-defined FB (e.g., 0.5 V) to understand how the resulting device degradation evolves with stress time.

Fig. 39 shows tunneling stress results on a 200 GHz *npn* device with same emitter geometry but different device layouts (CBE vs CBEBC). It was confirmed that the pre-stress RB I_B was comparable for the two devices to maintain a constant stress current. Results for three different 100 GHz *npn* devices with scaled emitter geometries and similar device layouts are also shown in Fig. 39. It was earlier confirmed that the RB EBJ I_B for these devices scaled according to their emitter area for each device generation. The 50 GHz devices did not show any observable change from the above stress, and so are not shown here in Fig. 39.

The results in Fig. 39 clearly indicate that there is some saturation of the I_B degradation at large stress times (not reported in [42]) that follows the initial power-law dependence. However, the 200 GHz devices show some variation in the amount of degradation at large stress times with a change in the device layout. Compared to the 200 GHz devices, the stress-time dependence of degradation for the 100 GHz devices is similar, but less in magnitude. The magnitude of the degradation clearly shows a direct correlation to the perimeter/area (P_E/A_E) ratio of the emitter geometry. This indicates that even if the areal component of the tunneling stress current is dominant for these devices, the damage to the EBJ related to trap-generation caused by tunneling stress is more dominant at the periphery of the emitter (localized to the EB spacer oxide-silicon interface). No previous studies have reported similar degradation effects resulting from the tunneling stress on EBJ of SiGe HBTs, and how that evolves with the device performance scaling.

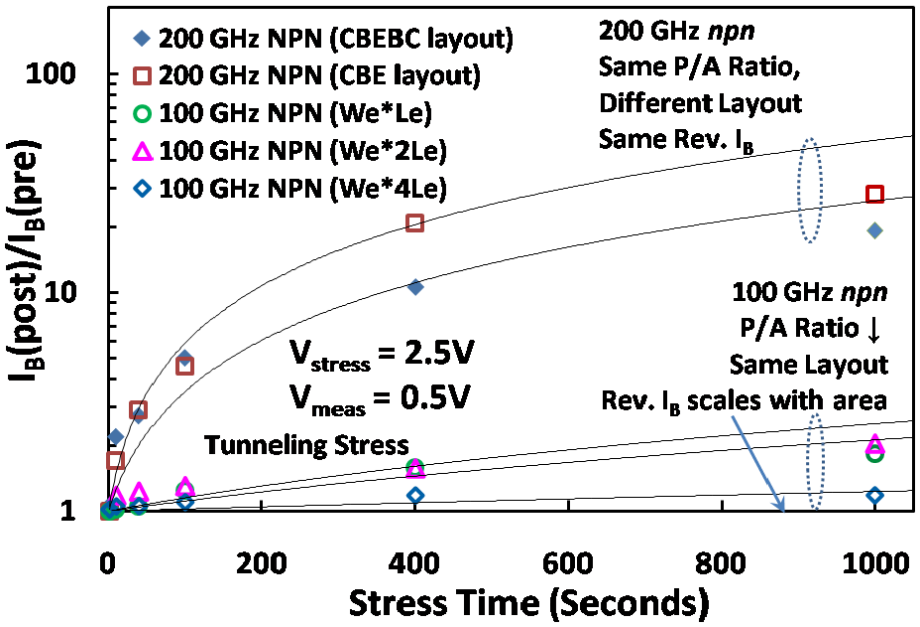

Figure 39: Comparison of the I_B degradation for 100 GHz and 200 GHz npn devices when subjected to RB EBJ stress in the tunneling current regime, including effects from device layout and emitter geometry variations [5].

6.6 Summary

This study has demonstrated for the first time that RB EBJ tunneling current for state-of-the-art SiGe HBTs can be well-modeled in a TCAD simulator with BTBT current. In addition, BTBT and TAT mechanisms can be incorporated judiciously in the simulator to model the RB EBJ I_B for multiple generations of SiGe HBTs (*npn* and *pnp*), and to reliably predict the trends that results from device scaling. Experimental data have been used to confirm that the simulation results are quite comparable in their scaling trends; within reasonable assumptions. Finally, accelerated tunneling stress results show that even though the tunneling current in these devices has a strong areal dependence, the damage from the stress is more localized at the emitter periphery of these devices.

CHAPTER 7

MIXED-MODE STRESS DEGRADATION MECHANISMS IN *PNP* SIGE HBTS

This work provides the first experimental investigation and evidence of the mixed-mode stress damage mechanisms (hot holes vs. hot electrons) in *pnp* SiGe HBTs. An increase of I_C resulting from MM stress (high stress-I_E and high V_{CB}) on *pnp* SiGe HBTs has been observed for the first time, and evidence of the physical mechanism causing this degradation is provided here through measurements and TCAD (technology computer-aided design) simulations [6].

7.1 Experimental and Simulation Details

The SiGe HBTs studied here are from a first-generation complementary SiGe BiCMOS process technology on thick-film SOI (silicon-on-insulator) [8, 18]. A representative cross-section of the *pnp* devices used in this study is shown in Fig. 40. The *pnp* and *npn* HBTs are designed for matched peak cut-off frequencies (f_T = 27 GHz at V_{CE} = 5 V) and open-base collector-emitter breakdown voltages (BV_{CEO} = 6.0 V). Two different emitter geometries of 0.4×0.8 μm^2 and 0.4×3.2 μm^2 are examined in this work.

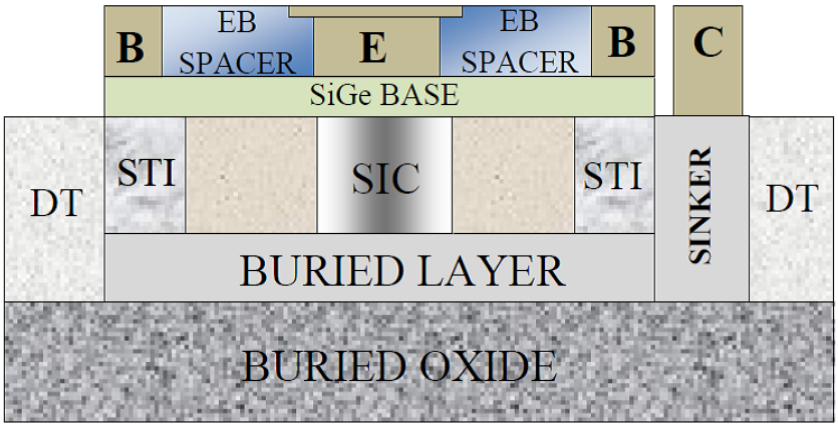

Figure 40: Schematic cross-section of a SiGe HBT from the complementary BiCMOS process (EB: emitter-base, STI: shallow trench isolation, DT: deep trench, E: emitter, B: base, C: collector, BURIED LAYER: sub-collector, SIC: selectively-implanted collector) [6].

The reliability of these complimentary SiGe HBTs was characterized with mixed-mode, reverse-EB (emitter-base) open collector (OC) and reverse-EB forward-biased collector (FC) stress. The M-M stress method uses a high reverse-biased collector-base (CB) junction with a forced constant emitter current (I_E) which drives I-I and collector current in the device to accelerate damage. Typical I_E values used correspond to the collector-current density (J_C) at peak f_T. The reverse EB stress involves reverse-biasing the EB junction beyond breakdown voltage with either an open-collector (OC) or forward-biased CB junction (FC) configuration. The forward and inverse Gummel characteristics are measured at fixed time intervals to observe the evolution of the device degradation with stress time. The changes in I_C and I_B are compared after 1000 seconds of stress time for different M-M stress conditions unless mentioned otherwise.

The avalanche multiplication factor (M-1) was extracted using the forced-I_E output characteristics method that is less sensitive to self-heating effects. In addition, this method allows for decoupling the effects of avalanche multiplication and Early effect on the increase of I_C with increasing V_{BC} [9]. All devices were measured at room temperature (RT). Device measurements were performed on-wafer using an Agilent 4155C Semiconductor Parameter Analyzer. Only standard devices available in the process technology were used for this investigation [18]. Two-dimensional (2-D) device simulations were performed using the Sentaurus Device simulator [56]. A 2-D device structure representative of the actual *pnp* HBT device being stressed was used for the simulations.

7.2 Results and Discussion

Fig. 41 shows the change of the forward Gummel characteristics of the *pnp* HBT with M-M stress. The M-M stress degradation mechanism is depicted in the inset of Fig. 41. An excess non-ideal base current (I_B) is observed at low-injection due to carrier recombination through traps generated by hot carriers (resulting from I-I in the CB junction) in the EB space-charge region at the Si/SiO_2 interface under the EB spacer. These traps act as G-R

(generation-recombination) centers to increase the SRH (Shockley-Read-Hall) recombination contributing to the non-ideal component of the base current [47, 48, 82]. Interestingly, however, an excess I_C is also observed in the low- to medium-injection regime. Fig. 42 shows that the corresponding current gain (β) decreases at low bias (due to the excess I_B) but also increases at moderate to strong bias with stress time. An increase in β has been previously reported for *pnp* Si bipolar junction transistors (BJTs) due to a decrease of I_B under high current stress [83]. As can be observed from Fig. 41, however, the I_B at medium- to high-injection is effectively unchanged and the increase in β is mainly attributed to an increase in I_C after the M-M stress.

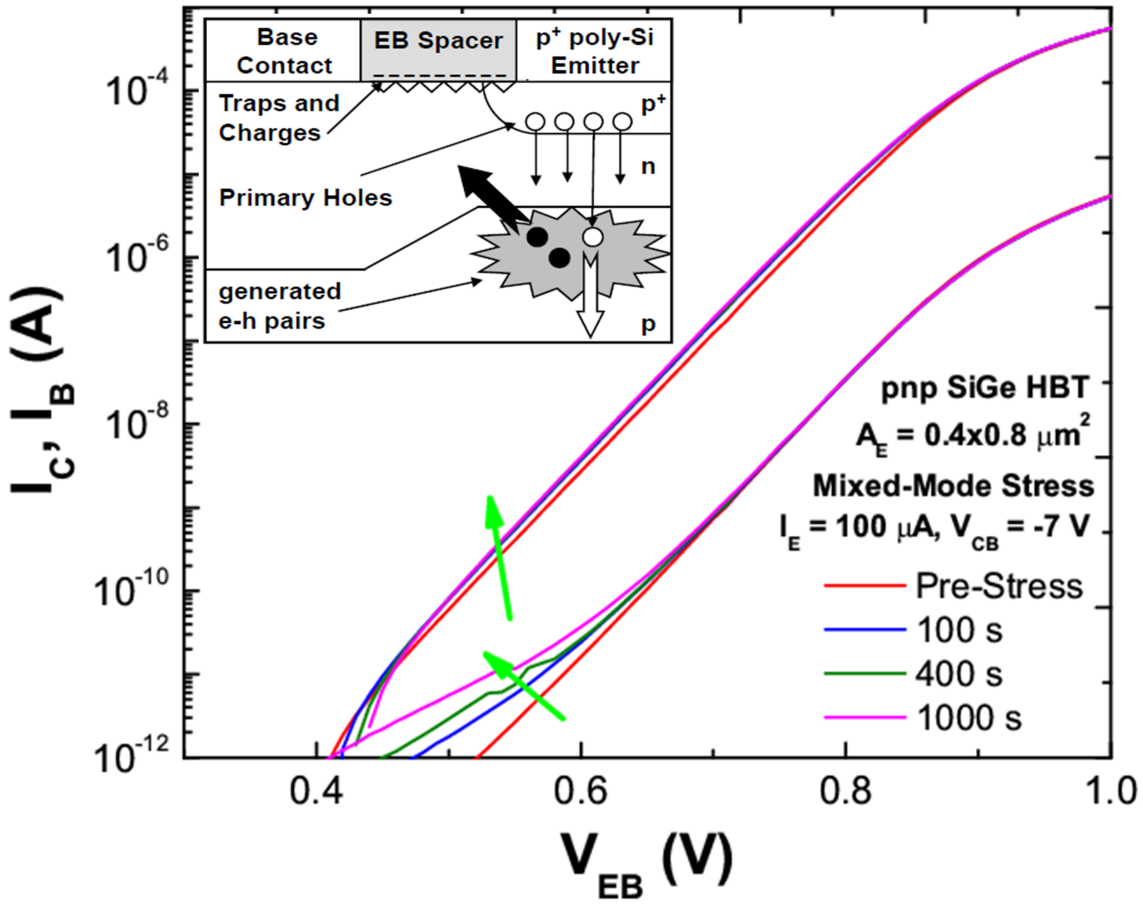

Figure 41: Gummel plots of a *pnp* SiGe HBT as a function of mixed-mode stress time. (Inset) Mixed-mode damage mechanism in a *pnp* SiGe HBT [6].

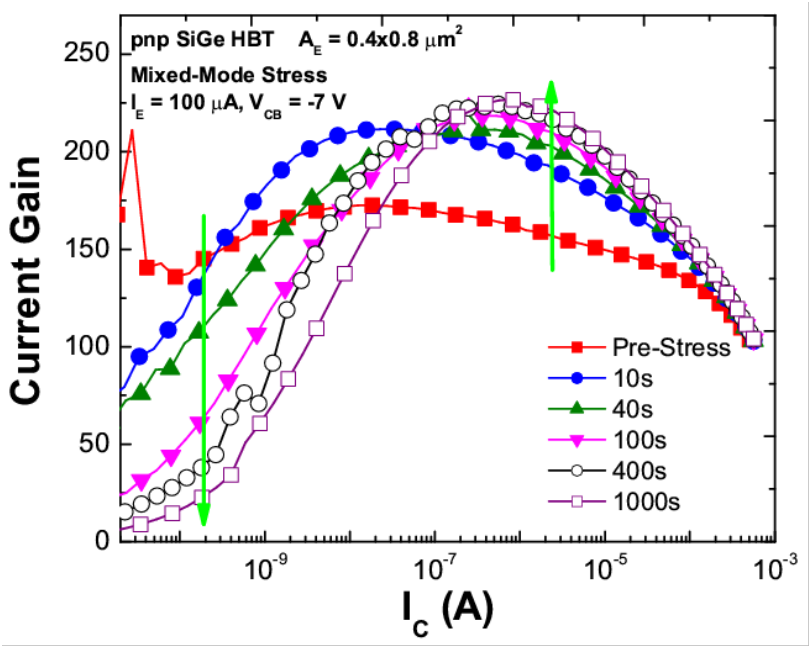

Figure 42: Current gain evolution of a *pnp* SiGe HBT as a function of the mixed-mode stress time [6].

To gain an insight into the underlying physical mechanism behind the increase of I_C with M-M stress, the ratios of the post- to pre-stress I_C (I_C ratio) for different M-M stress conditions were examined as a function of M-1. A typical plot of M-1 versus V_{CB} as a function of the stress I_E is shown in Fig. 43. The M-1 is a strong function of V_{CB} but decreases with increasing I_E due to high injection effects and self-heating. A typical plot of I_C ratio versus stress time as a function of I_E for a 0.4×0.8 μm^2 device stressed at a V_{BC} of 5.0 V is shown in Fig. 44. The I_C ratio increases linearly with logarithmic stress time (for any value of stress I_E) but the ratio decreases with increasing stress I_E. The decrease in I_C change with increasing stress I_E at the same V_{BC} correlates well to the decrease in M-1 with increasing stress I_E at the same V_{BC} (as seen in Figs. 43 and 44). A comprehensive plot of the I_C ratio versus M-1 extracted for the multiple M-M stress conditions and two device geometries is shown in Fig. 45. As can be observed, the change in I_C correlates well with the extracted M-1, suggesting an avalanche multiplication driven physical mechanism underlying the post M-M stress I_C behavior. Additionally, the device with an emitter geometry having a higher perimeter/area (P_E /A_E) ratio shows a larger

change in I_C at the same M-1, clearly indicating that the increase in I_C is associated with the emitter periphery of the transistor.

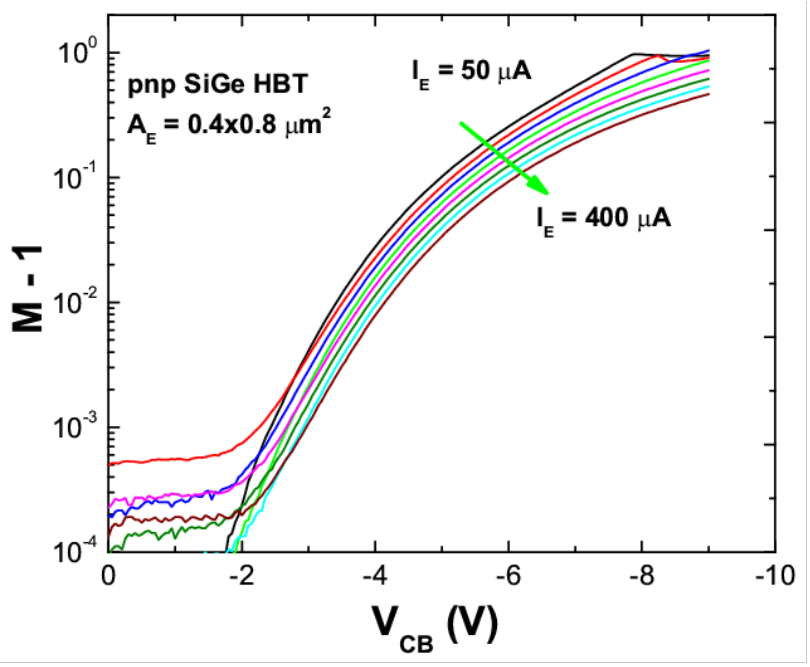

Figure 43: Avalanche multiplication factor M-1 measured for *pnp* SiGe HBTs [6].

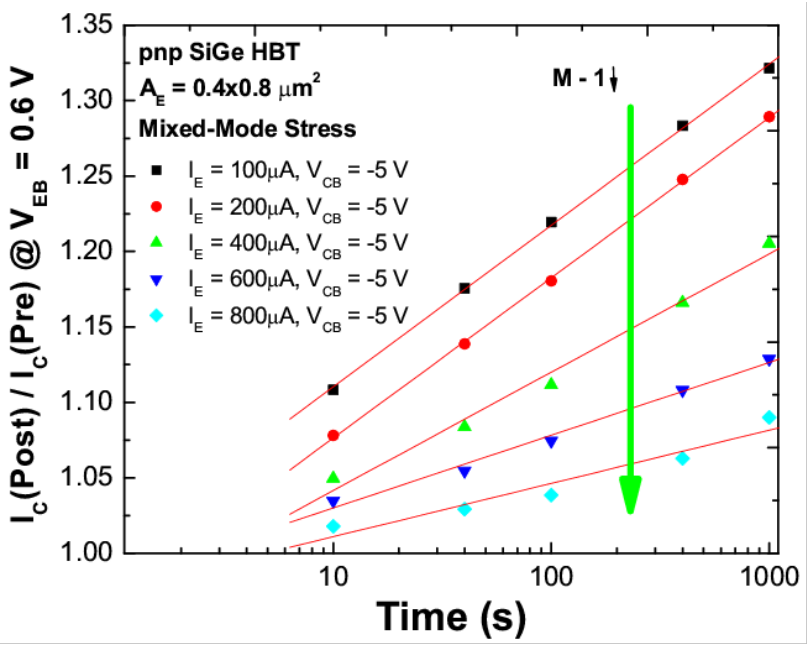

Figure 44: Change in collector current of the *pnp* SiGe HBTs with mixed-mode stress time for different stress conditions [6].

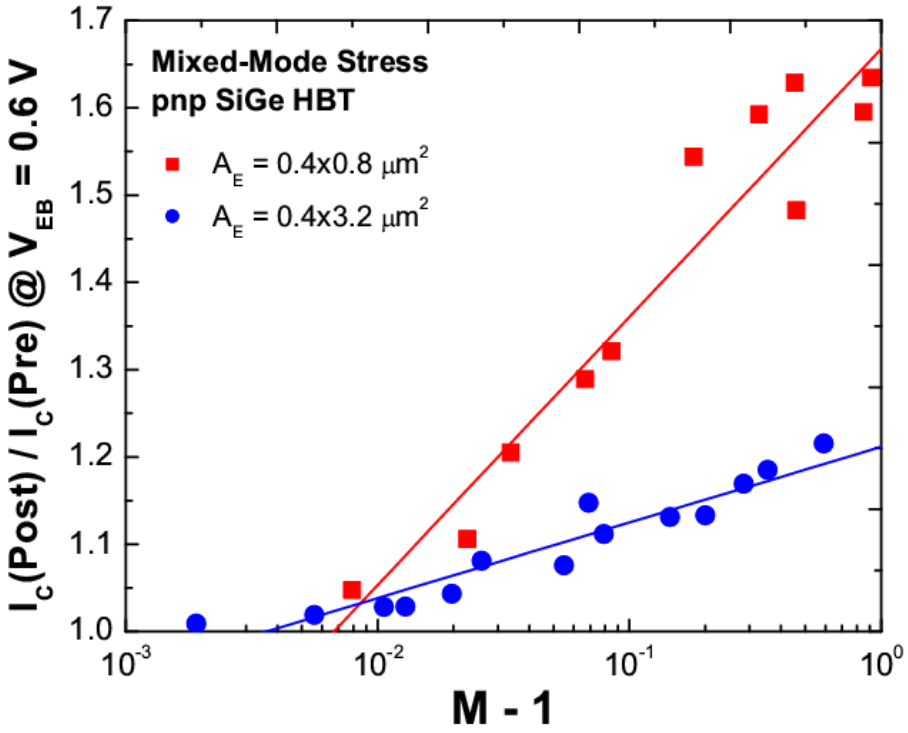

Figure 45: Change in I_C versus M-1 measured for different mixed-mode stress conditions on two *pnp* SiGe HBT device geometries [6].

Fig. 46 shows the output characteristics of the *pnp* device measured using the forced-I_E method. This data clearly shows evidence of "pinch-in" related voltage instabilities, though at V_{BC} values that are higher than the values used for M-M stress in this study, and only at the lowest current (I_E) levels. The onset of "pinch-in" is associated with the creation of localized "hot-spots" within the device, which affects the creation and distribution of hot-carriers, and hence the location of the damage resulting from the M-M stress. However, considering the measured output characteristics it can be concluded that there is no evidence of any "pinch-in" associated contribution to the mixed-mode stress damage observed for the *pnp* devices.

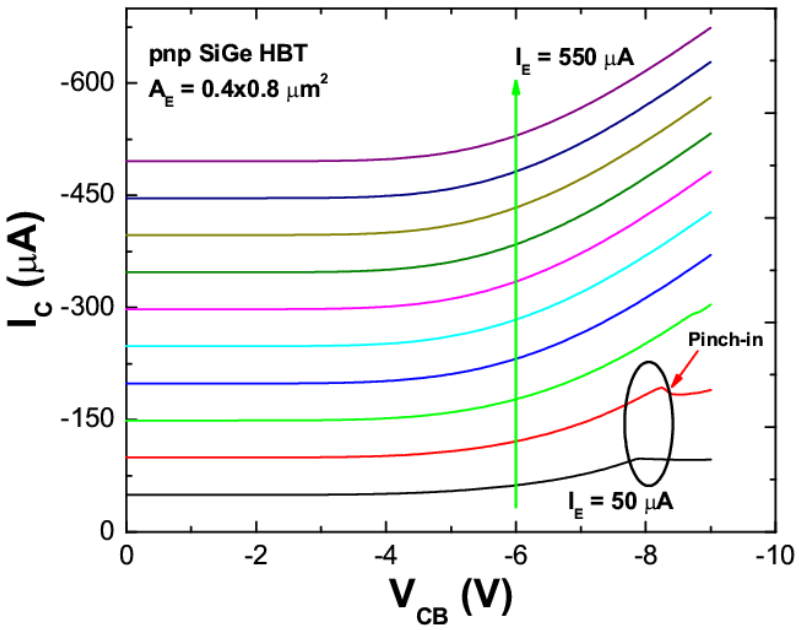

Figure 46: Forced-I_E output characteristics of the *pnp* SiGe HBT [6].

While, as argued above, the change in I_C can be attributed to hot-carriers generated in the CB junction, there is still uncertainty about whether hot holes or hot electrons are responsible for the damage. Previous simulation studies on the mixed-mode damage of *npn* SiGe HBTs using hydrodynamic simulations have suggested the importance of hot electrons [47], while Monte-Carlo simulations (of a different *npn* SiGe HBT technology) attributed the damage mainly to hot holes [48]. To identify the role of hot electrons versus hot holes in the M-M damage spectrum, the M-M stress on the examined *pnp* SiGe HBTs was followed by a reverse EB with FC stress on the same device. As can be seen from Fig. 47, the excess I_C decreases in value with increasing FC stress time. It is known that the FC stress damage is caused by hot holes (electrons) in *pnp* (*npn*) transistors [82]. Since the injection of hot holes into the EB spacer oxide decrease the excess I_C resulting from the preceding M-M stress, it is inferred here that this positive charge of the injected hot holes compensates the pre-existing negative charge in the spacer oxide formed predominantly by injection of hot electrons during the M-M stress. This compensation process decreases the I_C to its pre-M-M stress value, thus proving that the excess I_C during the M-M stress is

caused by embedded negative charges within the EB spacer oxide close to the oxide-silicon interface. This decrease in I_C with FC stress cannot be attributed to thermal annealing since the excess non-ideal I_B component continues to increase with the FC stress as shown in Fig. 47. The ratio of the post- to pre-stress I_B (I_B ratio) in Fig. 48 increases logarithmically with logarithm of stress time during the reverse EB with FC stress. A device with a higher P_E /A_E ratio shows a larger change in I_B ratio, clearly indicating that the excess non-ideal I_B is associated with the creation of traps under the EB spacer peripheral to the emitter geometry. Furthermore, an OC stress on the *pnp* device (which causes hot electron injection into the spacer) increases I_C (albeit to a lesser extent), thus providing additional evidence of the predominant role of hot electrons in the M-M stress damage of *pnp* HBTs. The data and experimental evidence presented here thus prove that changes observed in I_C and I_B ratios have their origins in different types of mixed-mode damage within the *pnp* device resulting during the same M-M stress.

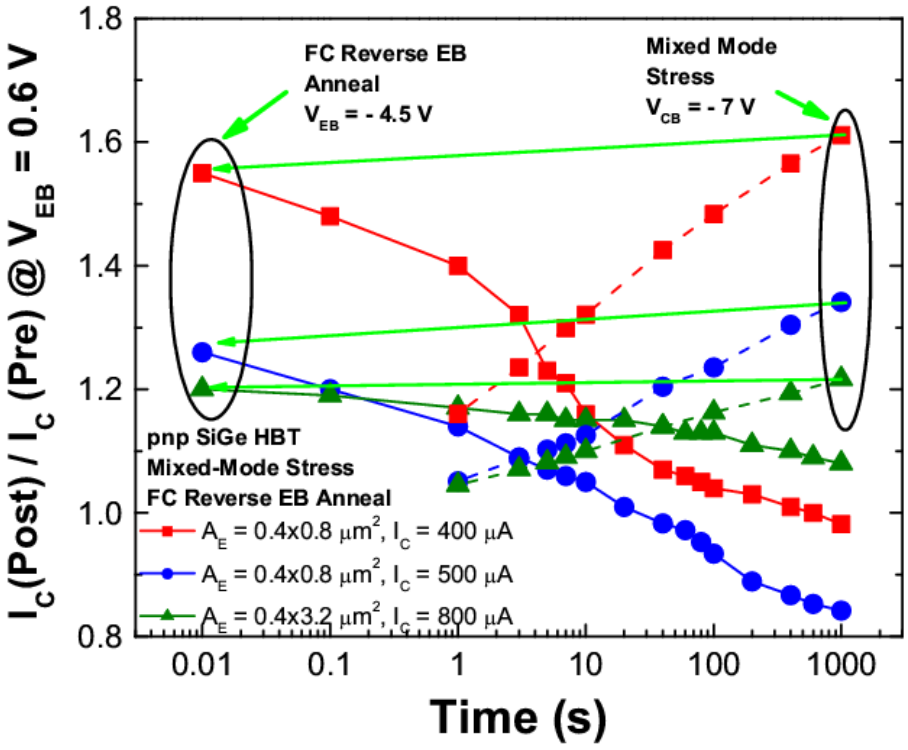

Figure 47: I_C ratio change versus mixed-mode stress time followed by reverse emitter-base with forward-collector stress time on the same *pnp* SiGe HBT [6].

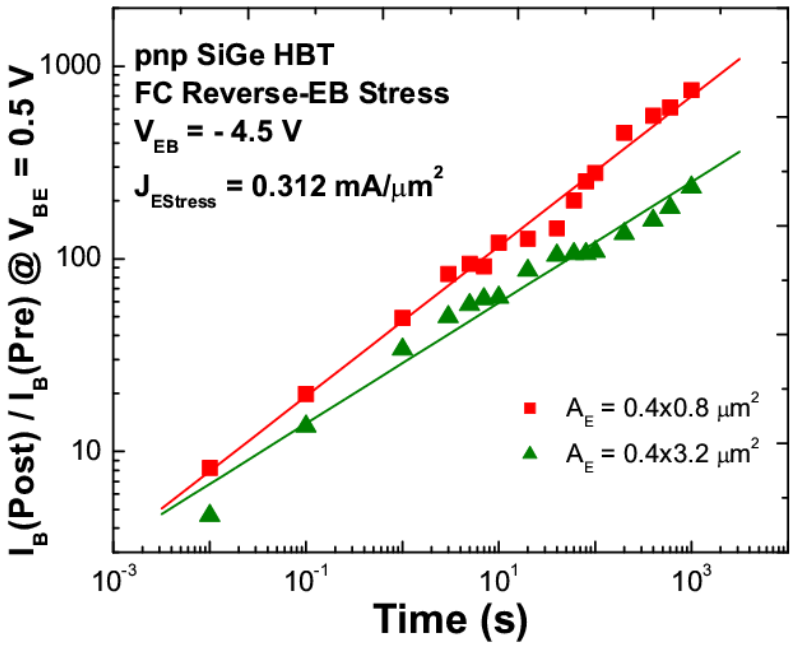

Figure 48: I_B ratio change with stress time for reverse emitter-base with forward-collector stress on *pnp* SiGe HBTs [6].

Based on the above experimental evidence, the physical damage mechanism in *pnp* HBTs during the M-M stress can be explained as follows. Energetic electrons (predominantly) generated by the I-I process at the BC junction during the M-M stress are transported across the base to reach the region under the EB spacer, where they create traps by breaking Si-H bonds at the interface, resulting in an excess non-ideal I_B. Some of these carriers have sufficient kinetic energy to be injected over the potential barrier (at the spacer/Si interface) into the EB spacer oxide layer, leading to a fixed charge density within the spacer near the Si interface. The fixed charge in the EB spacer acts to first deplete and then invert the base region under it; thus causing an increase in the effective emitter area and leading to an increase in I_C. A similar increase in the post-radiation I_C has been reported earlier for *npn* BJTs [12].

TCAD simulations show that for a high enough negative charge density near the EB spacer/Si interface and low enough doping under the spacer, the surface of the n-type extrinsic base can first be depleted and then inverted (p-type) during the M-M stress. When the hole concentration in this inversion layer becomes comparable to that at the edge of

the p-type emitter, the inversion layer forms an extension of the physical emitter at the periphery, increasing the effective emitter area (A_E) and thus leading to a higher I_C (Fig. 49 and inset). Furthermore, as the EB junction is forward-biased, holes are injected into the n-type base from both the p-type emitter and the p-type inversion layer which functions as an extension of the emitter area, thus leading to a higher I_C. For a surface doping of 5×10^{15} cm^{-3}, a charge accumulation of 5×10^{11} cm^{-2} is shown to be sufficient to cause a significant change in the collector-current of the device (Fig. 49).

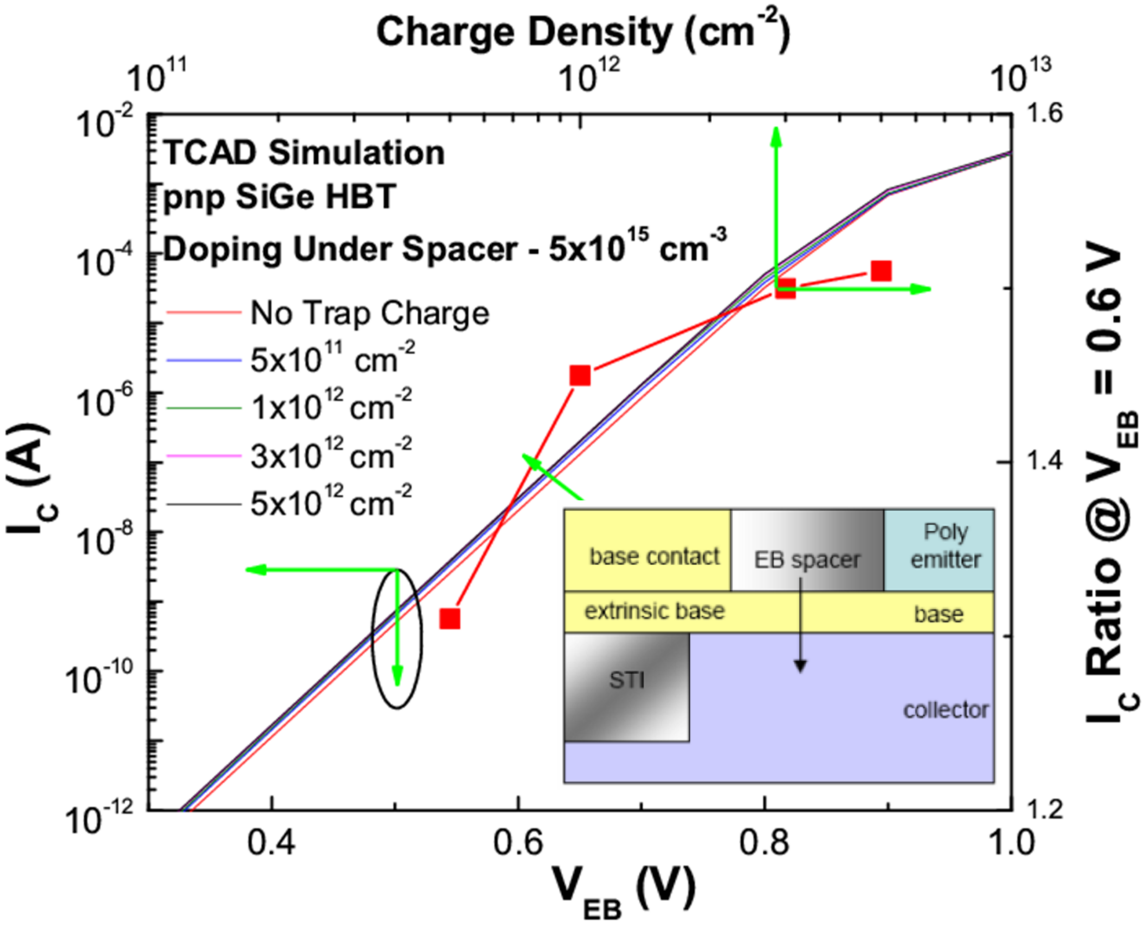

Figure 49: TCAD simulations of the I_C and I_C ratio change for a *pnp* SiGe HBT at different charge densities within the emitter-base spacer. (Inset) Device structure used for TCAD simulations [6].

It is clearly shown experimentally for the first time what type of hot carriers actually drives the mixed-mode damage process. Furthermore, TCAD simulations with different doping concentrations and thicknesses of the low-doped region under the EB spacer/Si

interface (Fig. 50) seem to indicate that inversion of the extrinsic base is possible for a relatively high doping of $10^{17}\ cm^{-3}$ and a thickness of only 2 nm, and this leads to a similar order of magnitude of I_C change as observed from the M-M stress data. The TCAD simulations assumed a charge/trap density of about $5 \times 10^{12}\ cm^{-2}$, a value typically applied to these kinds of studies.

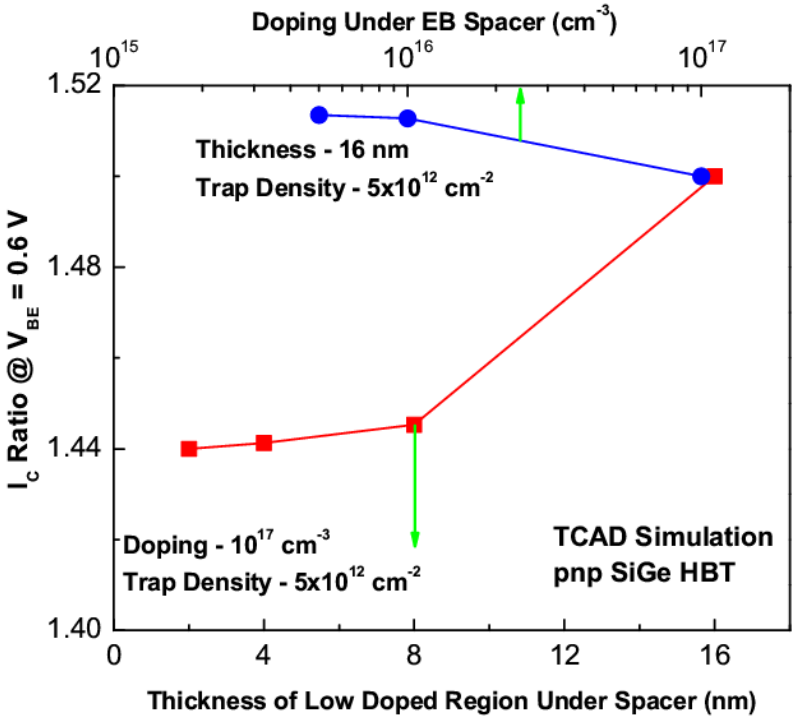

Figure 50: TCAD simulations of I_C ratio change of the *pnp* SiGe HBT for different doping concentration and thickness of the low-doped layer under the emitter-base spacer [6].

No perceptible I_C change was noticed for *npn* HBTs from the same complementary SiGe BiCMOS process technology when subjected to similar mixed-mode stress conditions. Increase in non-ideal I_B was observed, similar to what has been reported by earlier studies [47, 48, 82]. This difference in post-stress I_C behavior for *npn* versus *pnp* devices was observed even though the M-1 values for the *npn* devices being stressed were found to be comparable to those of the *pnp* devices stressed in this study, as is evident from comparing Fig. 45 and Fig. 51. This difference in the M-M stress response of *npn* and *pnp* devices from the same process technology can be attributed to inherent differences of the two devices.

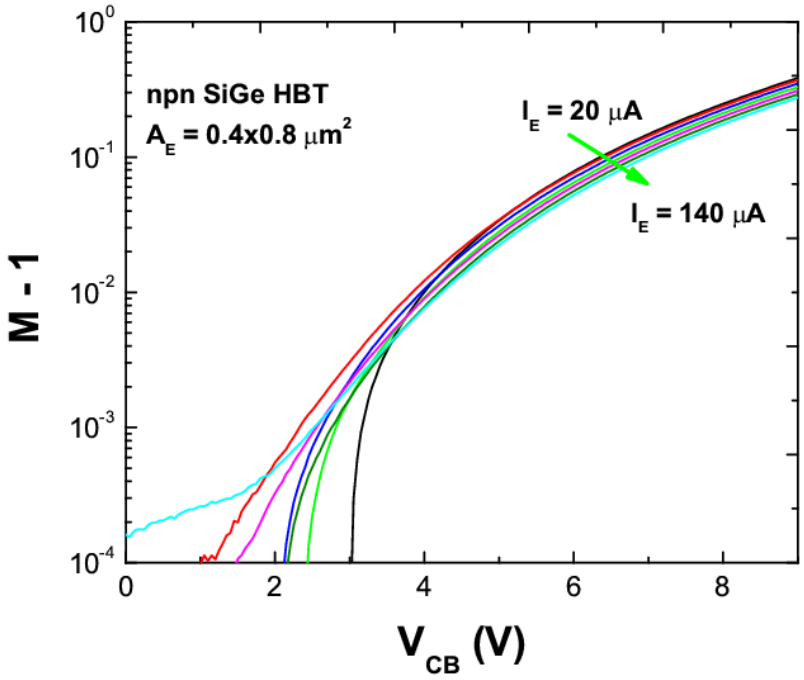

Figure 51: Avalanche multiplication factor M-1 measured for a *npn* SiGe HBT available from the same process technology [6].

7.3 Summary and Conclusion

Experimental evidence for the type of hot-carriers causing device degradation in *pnp* SiGe HBTs during M-M stress is provided for the first time. Comprehensive stress results presented as a part of this study with different device geometries and stress conditions prove that the kinds of mixed-mode damage resulting in changes of I_C and I_B are different. Detailed TCAD simulations show that typical charge densities in the EB spacer are enough to create an I_C change even for a very narrow moderately doped region of the extrinsic base under the spacer/Si interface. It can be concluded that a difference in mixed-mode reliability for *npn* and *pnp* HBT devices in a complementary SiGe BiCMOS process will be an important challenge for robust circuit design, and needs to be investigated in further detail both at a fundamental and an application level.

CHAPTER 8

CRYOGENIC PERFORMANCE OF THE BEST-OF-BREED *NPN* AND *PNP* SIGE HBTS

This study leverages cryogenic performance enhancement of SiGe HBTs using temperature as a scaling-lever to extend the maximum achievable performance of best-of-breed *npn* and *pnp* SiGe HBTs, and to compare their cryogenic performances as part of a viable SiGe BiCMOS technology.

The work is presented in two parts. In the first part, a record cryogenic performance of a best-of-breed *npn* SiGe HBT as part of an *npn* only SiGe BiCMOS process is studied and presented [7]. In the second part, *npn* and *pnp* SiGe HBTs with comparable performance from a best-of-breed C-SiGe BiCMOS technology that has the highest reported *pnp* SiGe HBT performance is studied, and results are presented with analysis for the *npn* and *pnp* devices with comparable performance, as well as between the best-of-breed *npn* and *pnp* devices. In a nutshell, this work provides the first cryogenic performance scaling studies of *pnp* SiGe HBTs as part of a C-SiGe HBT platform.

8.1 Device Technology Details

The best-of-breed *npn* SiGe HBTs used for this investigation are from a commercially-available 130 nm double-polysilicon *npn*-only 4th generation SiGe BiCMOS platform (IHP G2). This SiGe technology is based upon those presented in [84] and [85], but it does differ in some aspects in the fabrication flow, resulting in slightly lower performance. The f_{MAX} enhancement primarily comes from the lateral and vertical scaling and a reduced thermal budget, together with changes in the emitter-base composition, collector formation and salicidation. Some of the key device design aspects from [84] were discussed in [86]. Since simultaneous f_T and f_{MAX} scaling necessarily involves optimization of the device parasitics as well as vertical profile scaling, and as power gain is a more important device metric than

current gain for most high-frequency circuits, achieving $f_{MAX} \geq f_T \geq f_{MAX}/2$ continues to be a desirable target for SiGe HBT scaling as it moves towards THz speeds [87]. Important device design and process technology challenges to scale this device technology to THz performance at 300 K have been highlighted in [86].

The best-of-breed *pnp* SiGe HBTs are part of the highest performance C-SiGe technology. It is a 250 nm second generation complementary SiGe BiCMOS platform with shallow-trench isolation and implants for the collector formation, as well as implants for the vertical isolation of the *pnp* collector from the p- substrate (IHP SG25H3P). The *npn* and *pnp* devices were optimized for comparable performances [19, 21, 88, 89].

8.2 Measurement Details

The *ac* performance of the SiGe HBTs was measured down to 4.3 K on a custom-designed, on-wafer, open-cycle liquid helium cryogenic probe station. SOLT and TRL calibration were performed with an Impedance Standard Substrate (ISS) to remove the cable and probe losses at each temperature. The results obtained using both calibration techniques were comparable. S-parameters were measured to 50 GHz using an Agilent E8364C PNA. On-wafer "Open" and "Short" calibration structures located adjacent to the device were used to deembed the pad parasitics (i.e., those extrinsic to the device) from the measured S-parameters. Additional *dc* characterization was performed on packaged die down to 13 K using a closed-cycle cryogenic system. The results presented here were reproducible over multiple measurements and samples.

8.3 Results from Characterization of best-of-breed *npn* HBTs

The SiGe HBTs showed reasonably ideal *dc* Gummel characteristics over temperature, all the way down to 4.3 K. The base and collector current idealities were similar to those reported in earlier studies [49, 50]. The *dc* current gain (β_{DC}) and transconductance (g_m) monotonically increase with cooling, as expected, which is evident from the plots of peak

β_{DC} and peak g_m shown in Fig. 52(a). Peak g_m rises with cooling from 161 mS at 300 K, and the peak β_{DC} increases to 3,369 at 78 K and then stays nearly constant to 4 K. Carrier freeze-out in the device is insignificant in these devices due to the heavy doping (above the Mott-transition) in the emitter, base and collector regions. In Fig. 52(b), the packaged devices show a decrease in the measured thermal resistance (R_{TH}) from 1,891 K/W at 300 K to 1,450 K/W at 78 K. R_{TH} typically increases below 78 K due to a combination of deep cryogenic scattering mechanisms and package-induced measurement limitations, as discussed in [90, 91]. The low R_{TH} values achieved in these SiGe HBTs at 300 K can be mainly attributed to the device structure, key elements of which are absence of deep trench isolation, and careful layout optimization, which in turn helps to improve the *ac* performance and reduce electro-thermal instabilities within the device [84–86]. The avalanche multiplication coefficient (M-1), measured at low-injection (to decouple it from any significant self-heating effects), provides a qualitative read on the anticipated reliability of the device from impact-ionization driven damage mechanisms originating in the collector-base junction (e.g., mixed-mode electrical stress). The low-injection M-1 values extracted using a common-base forced-I_E measurement are shown in Fig. 52(b), and increase only minimally from 300 K to 78 K, and are then nearly constant to 4 K.

Fig. 53(a) shows near ideal (~20 dB/dec roll-off) small-signal current gain (h_{21}) up to 50 GHz. These values were extracted from the de-embedded S-parameters at 300 K and 4.3 K, for a device biased under peak f_{MAX} bias conditions at those temperatures. The 4.3 K data were restricted to above 8 GHz to reduce the measurement time, thus minimizing cryogen consumption. The plot of $MUG^{1/2}$ (Masons unilateral gain) under similar bias conditions also shows near ideal behavior (~20 dB/dec roll-off) up to nearly 50 GHz, at both 300 K and 4.3 K, confirming the robustness of the measurements. The f_T and f_{MAX} values were then extracted by fitting -20 dB/dec lines through the near-ideal region of the measured h_{21} and $MUG^{1/2}$ values, respectively, and extrapolating them to unity gain. This technique of fitting and extrapolation provides a more robust error-estimation method for

f_T/f_{MAX} extraction compared to extrapolating from a specific frequency point in the near-ideal region, thereby introducing a smaller margin for error. Fig 53(b) shows f_T/f_{MAX} for the device as a function of collector current density at 300 K, 78 K and 4.3 K. At each temperature, the peak f_T/f_{MAX} values occur at comparable J_C, with the J_C at peak f_T increasing slightly with cooling due to an increase in the electron saturation velocity [9].

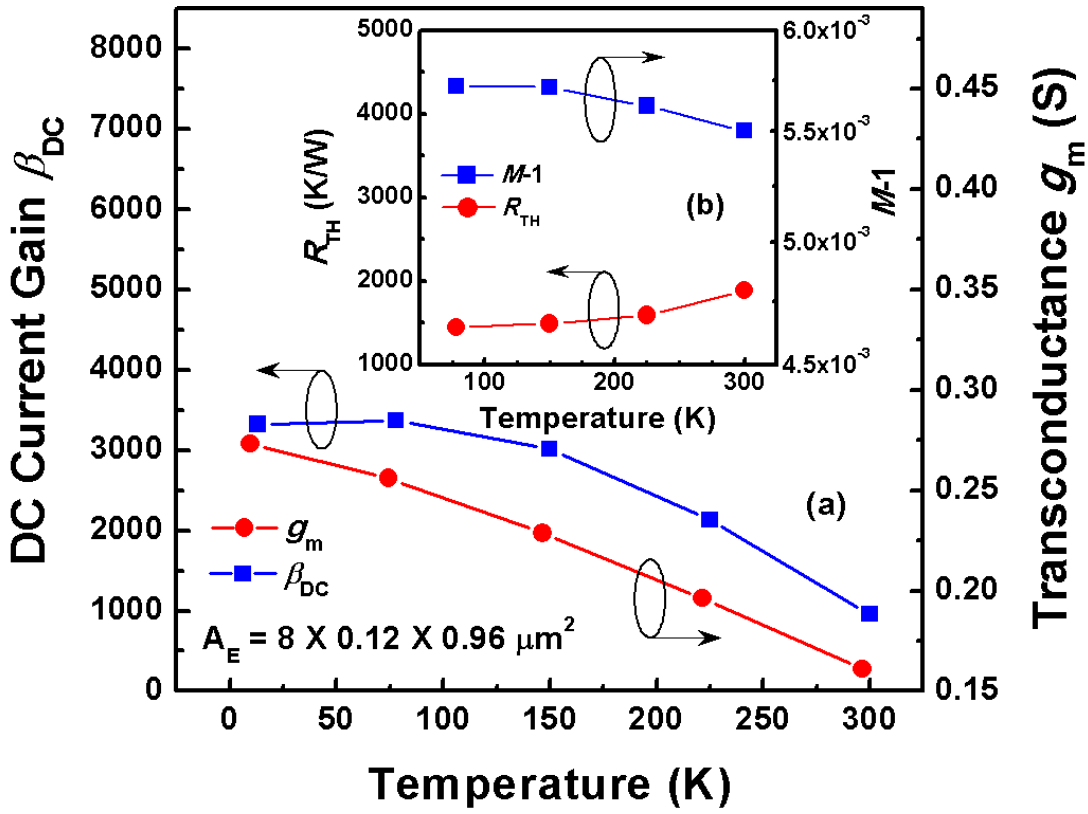

Figure 52: (a) Plots of peak transconductance and *dc* current gain measured as a function of temperature. (b) Plots of measured thermal resistance (R_{TH}) and avalanche multiplication coefficient (M-1, at V_{CB} = 1.2 V, $J_E = 10^{-5}A/\mu m^2$) [7].

The *ac* performance of this SiGe HBT is summarized in Table 5, which shows that both peak f_T and peak f_{MAX} increase with cooling down to 4.3 K. The SiGe HBT achieves a record peak f_{MAX} of 798 GHz at 4.3 K (417 GHz at 300 K). The extracted r_{bb} (using the impedance-circle technique described in [9]) remains fairly stable between 300 K and 78 K, and then decreases to 10.7 Ω at 4.3 K. The decrease in r_{bb} is mainly driven by the increase of carrier mobility inside the neutral base at 4.3 K, which is fully ionized since it is doped above the Mott-transition [91]. The forward transit time (emitter-to-collector delay - τ_{EC})

extracted from the extrapolated y-intercept of $1/2\pi f_T$ vs. $1/J_C$ plot shows that τ_{EC} decreases to 0.30 ps at 4.3 K [9].

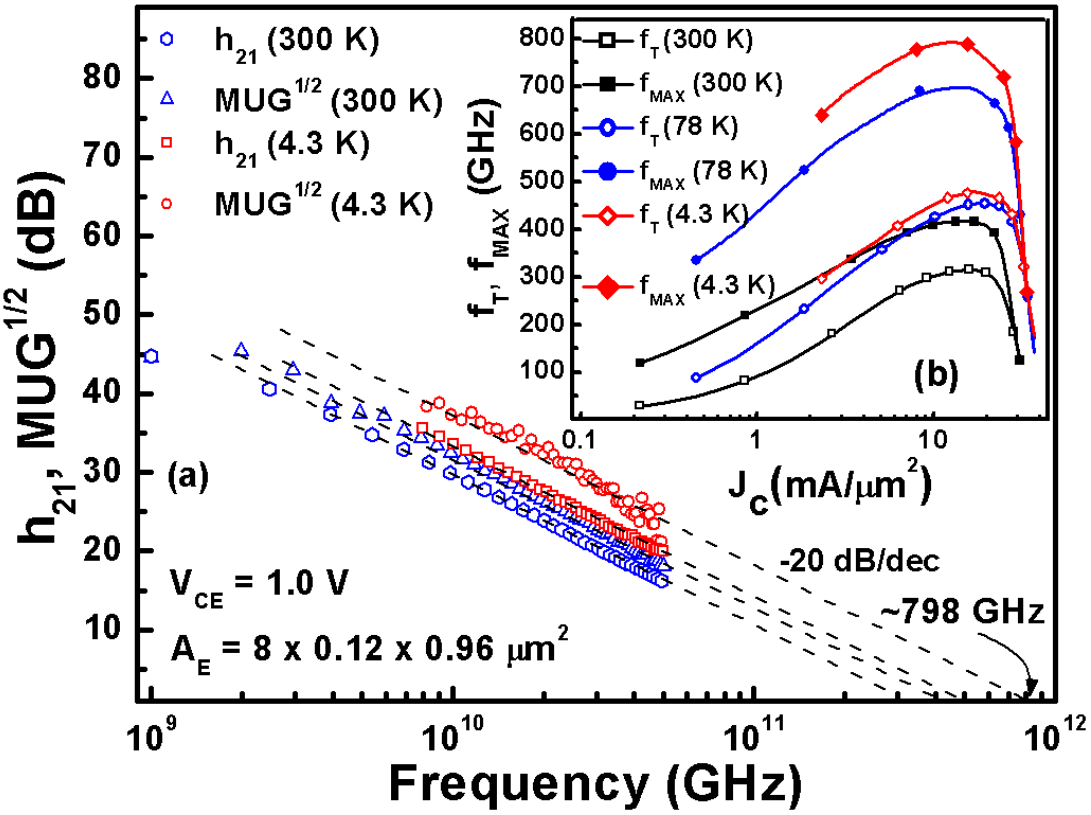

Figure 53: (a) Measured small-signal current gain (h_{21}) and Masons unilateral power gain ($MUG^{1/2}$) as a function of frequency. (b) Extracted unity gain cutoff frequency (f_T) and maximum oscillation frequency (f_{MAX}) as a function of collector-current density (J_C) [7].

The BV_{CEO} for the device (measured from the base-current reversal point using a standard common-base, forced-I_E technique) does not degrade significantly with cooling, remaining above 1.6 V at 4.3 K, thereby producing a record $f_T \times BV_{CEO}$ product of 800 GHz-V and a record $f_{MAX} \times BV_{CEO}$ product of 1,333 GHz-V at 4.3 K. It should be noted that even though we measured the S-parameters at a constant V_{CE} of 1.0 V to avoid any degradation down to cryogenic temperatures, it is clear from *dc* measurements that this device does not show any abnormal avalanche-multiplication and self-heating effects at 1.0 V $\leq V_{CE} \leq$ BV_{CEO}, and is thereby capable of being biased at a higher voltage, which can potentially result in an even higher peak f_T and f_{MAX}.

Table 5: Cryogenic Performance Summary of the *npn* SiGe HBT [7]

Figure-of-Merit	300 K	78 K	4.3 K
Peak f_T (GHz)	315	456	479
Peak f_{MAX} (GHz)	417	702	798
r_{bb} at Peak f_T (Ω)	18.1	18.9	10.7
Total Transit-Time τ_{EC} (ps)	0.39	0.31	0.30
J_C at Peak f_T (mA/μm^2)	16.2	17.1	18.1
Minimum BV_{CEO} (V)	1.63	1.65	1.67
$f_T \times BV_{CEO}$ (GHz-V)	513	752	800
$f_{MAX} \times BV_{CEO}$ (GHz-V)	680	1,158	1,333
$f_T + f_{MAX}$ (GHz)	732	1,158	1,277

Fig. 54(a) shows previously reported SiGe HBTs from different technologies (with a range of performance) measured down to cryogenic temperatures, and is compared with results from the present work. The data are plotted as peak f_{MAX} vs. peak f_T, and peak f_T vs. minimum BV_{CEO} (in Fig 54(b)). The devices shown were chosen to represent performance trends clearly without redundancy. The devices for which $f_{MAX} \geq f_T \geq f_{MAX}/2$ is satisfied, along with achieving a maximum for $f_T \times BV_{CEO}$ and aggregate performance ($f_T + f_{MAX}$) have the most desirable composite performance for optimized scaling of these devices to THz speeds. It is evident from Table 5 and Fig. 54 that the *npn* devices studied in this inverstigation satisfy that criteria. From the data shown in Fig. 54(a), it can also be clearly observed that the trends seen in the cryogenic performance enhancements are comparable to conventional device performance scaling at 300 K.

8.4 Results from Characterization of best-of-breed Complementary SiGe HBTs

The complementary SiGe HBTs (*npn* and *pnp*) showed fairly ideal *dc* Gummel characteristics over temperature, all the way down to 78 K for the *npn* and 4.3 K for the *pnp*. The base

and collector current idealities were comparable to that of *npn* only technologies. As expected and shown in Fig. 55, the peak transconductance/area increases monotonically with cooling (for both devices), while BV_{CEO} decreases monotonically with cooling (for both devices). The peak *dc* current gain (β_{DC}) also increases monotonically for both devices with decreasing temperature (not shown here). However the BV_{CEO} of the *pnp* is lower than the *npn* at 78 K and below due to a sharper increase in the current gain of the *pnp* compared to that of the *npn*.

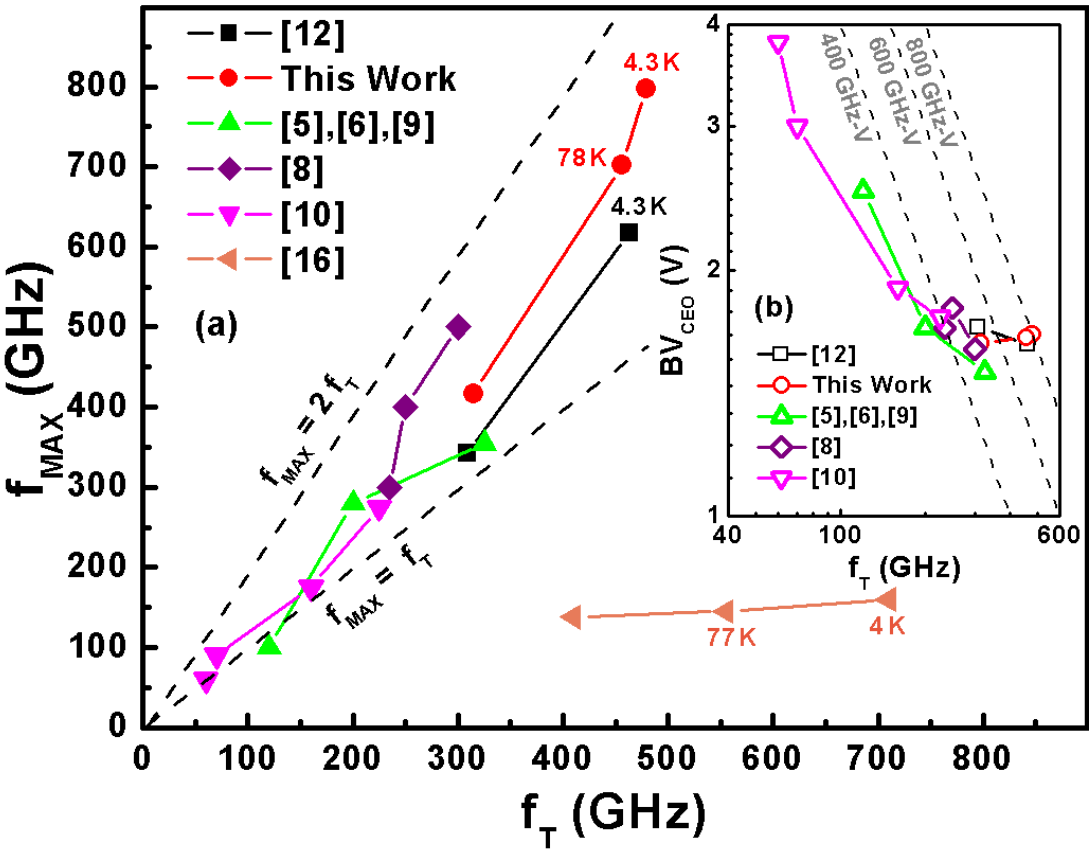

Figure 54: (a) Plot of f_{MAX} vs. f_T for reported data over temperature compared to the present work, showing the $f_{MAX} \geq f_T \geq f_{MAX}/2$ region between the two dashed lines, which offers the most desirable target for device scaling. All data presented are at 300 K unless mentioned with a temperature label next to the data point. (b) Plot of f_T vs. BV_{CEO} for reported SiGe HBT performance over temperature, and is compared to the present work, along with constant $f_T \times BV_{CEO}$ loci. Temperature is derived by correlating to the f_T values in (a). The reference numbers quoted correspond to those cited in [7].

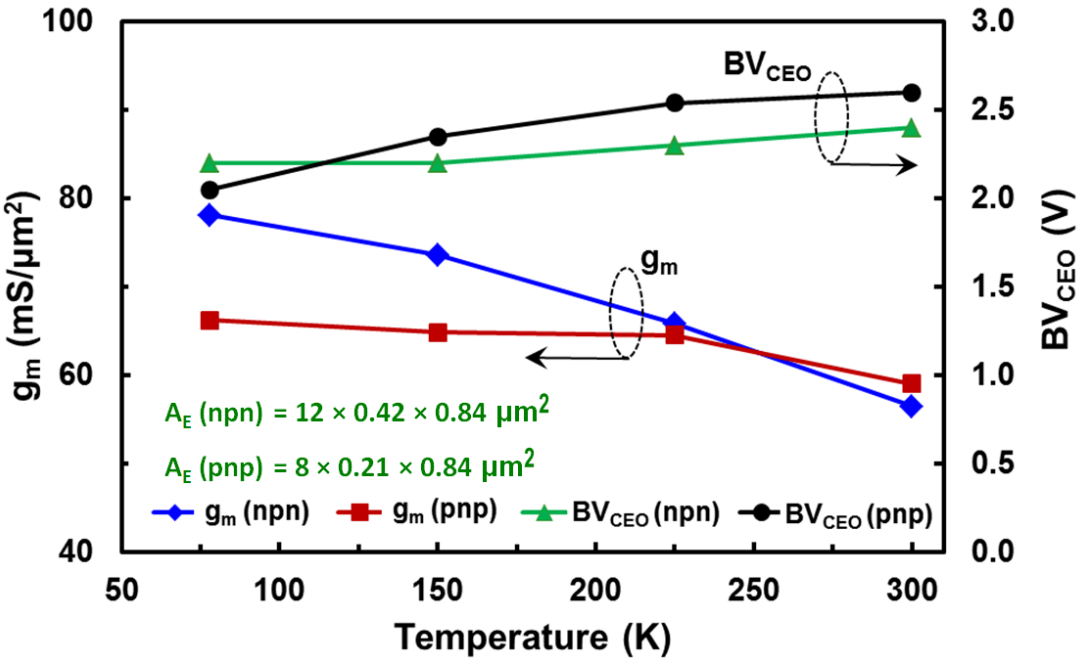

Figure 55: Plots of peak transconductance per unit area and BV_{CEO} measured as a function of temperature for the *npn* and *pnp* devices from the C-SiGe technology. Devices were measured down to 78 K

The complementary SiGe HBTs (*npn* and *pnp*) showed fairly normal *ac* characteristics over temperature , all the way down to 78 K for the *npn* and 4.3 K for the *pnp*. Extracted unity gain cutoff frequency (f_T) and maximum oscillation frequency (f_{MAX}) as a function of collector-current density (J_C) showed behavior very typical of existing SiGe HBTs (from both npn-only and C-SiGe technologies). Both small-signal current gain (h_{21}) and Mason's Unilateral Power Gain (MUG) extracted from measurements showed near-ideal roll-off with frequency for both devices down to the lowest cryogenic temperatures used for the measurements.

As shown in Fig. 56(a), both peak f_T and f_{MAX} of both devices increases monotonically with cooling to the lowest temperature used for the measurements. The *ac* performance matching of the *npn* and *pnp* devices worsens with cooling (while comparing peak f_T and f_{MAX}). The *pnp* shows a cryogenic ac performance of $f_{MAX} > 200$ GHz at 4.3 K, and this is the highest reported ac cryogenic performance of *pnp* SiGe HBTs till date. This is also the highest performance reported by any *pnp* device as part of a C-SiGe HBT technology.

When the f_{MAX} vs. f_T for both the *npn* and *pnp* devices at the different measurement temperatures are plotted in Fig. 56(b), it clearly shows that the device performance stays in the desirable range of operation for all measurement temperatures considered. This region is also highly desirable for suitable high-frequency power amplification properties of the device.

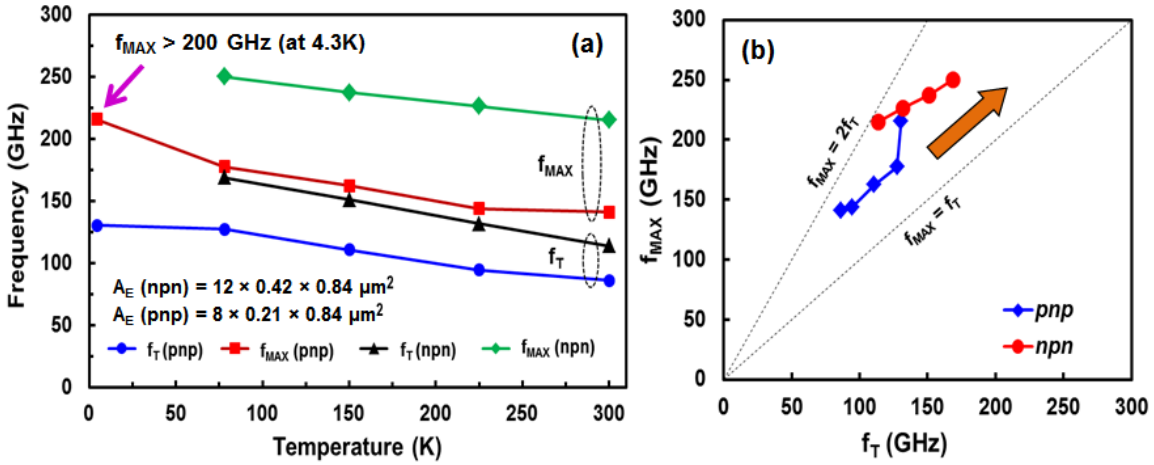

Figure 56: (a) Plots of peak f_T and f_{MAX} for both the *npn* and *pnp* devices as function of temperature. Only *ac* data for the *pnp* available down to 4.3 K. (b) Plots of peak f_{MAX} vs. f_T for both the *npn* and *pnp* devices at the different measurement temperatures. It delineates the $f_{MAX} \geq f_T \geq f_{MAX}/2$ region between the two dashed lines, which offers the most desirable target for device performance scaling.

Taken together, the cryogenic results for the complementary SiGe HBT devices imply that C-SiGe devices shows similar cryogenic performance enhancement like *npn*-only technologies. However, the C-SiGe devices might show worse performance matching at the lowest cryogenic temperatures compared to 300 K. Devices from C-SiGe technologies are expected to scale very similarly to the *npn*-only technologies as is evident from the cryogenic performance enhancements. It would be possible to evolve the performance of the C-SiGe devices while scaling very similar to that of the *npn* devices, so that the C-SiGe devices continue to maintain desirable high-frequency properties when scaled.

8.5 Comparison and Analysis of the *npn* and *pnp* cryogenic performance

Comparing the cryogenic characterization results of the *npn* and *pnp* devices from *npn*-only and C-SiGe technologies provide us with a few key insights.

The "best-of-breed" *pnp* devices are expected to maintain similar scaling trends as the "best-of-breed" *npn* devices, based on results evident from the cryogenic characterization and performance enhancements of these devices. While the *npn* performance is expected to exceed 0.8 THz soon at 300 K as part of a *npn*-only technology , the *pnp* performance is expected to exceed 200 GHz (for both f_T and f_{MAX}) soon as part of a C-SiGe HBT technology. Hence C-SiGe HBT technologies are expected to scale up to 200 GHz or beyond, as npn-only technologies gets slowly scaled close to THz performance.

8.6 Summary

A record peak f_{MAX} of 798 GHz at 4.3 K for a npn SiGe HBTs in a 130 nm SiGe BiCMOS technology platform has been demonstrated. Both peak f_T and f_{MAX} were found to increase with cooling, while BV_{CEO} does not show any significant degradation over temperature. The excellent thermal resistance and avalanche multiplication coefficient of these scaled SiGe HBTs has significant implications, suggesting that they should have attractive electro-thermal properties and reliability over temperature.

Second part of the study reported a record peak $f_{MAX} > 200$ GHz for a *pnp* at 4.3 K as part of a C-SiGe HBT technology. It is also the highest performance reported till date for a *pnp* as part of a C-SiGe HBT technology. The *pnp* devices showed cryogenic performance enhancement trends which are very similar to that of the *npn* devices. This is the first reported cryogenic characterization and performance scaling study of C-SiGe HBT devices. Overall, both of these studies taken together have significant implications for performance scaling of the *npn*-only and C-SiGe device technologies.

CHAPTER 9

SAFE-OPERATING-AREA CONSTRUCTION FOR HOT-CARRIER RELIABILITY AND THE ANNEALING OF HOT CARRIER INDUCED DAMAGE

This first part of this study leverages hot-carrier induced device degradation from mixed-mode stress on SiGe HBTs to provide new methods of constructing the SOA and comparing the reliability of two devices that have a comparable performance, like between the *npn* and *pnp* SiGe HBTs in a C-SiGe BiCMOS technology. In the second part of this study, new ways of characterizing the damage and annealing process in SiGe HBTs are proposed to provide an improved understanding of the physical mechanisms involved. Electrical stress methods are used to provide insight into the annealing behavior of hot-carrier induced damage, and to investigate the reaction-diffusion mechanism used to model hot-carrier damage (and annealing) in SiGe HBTs. Further, the investigation presented here is an experimental effort to characterize and develop understanding about the nature of the annealing processes across the output plane, while enabling improvement of the existing physics-based TCAD approach mentioned earlier in Chapter 2 with predictive models for annealing [results to be submitted for publication].

9.1 Device Technology

The devices used for the first part of the study were derived from a first-generation (with peak $f_T < 50$ GHz) complementary SiGe BiCMOS process technology with *npn* and *pnp* devices of comparable performance. Second part of the study leverages this technology as well as a *npn*-only third-generation SiGe BiCMOS technology (with peak $f_T > 200$ GHz) [92]. Both of these are bulk Si based process technologies. Two different types of device technologies from different generations of performance were leveraged to provide improved understanding of the concepts involved in this study, and how these concepts change as a function of performance scaling.

9.2 Stress and Characterization Methods

The first part of the study uses time-dependent device degradation from mixed-mode stress [46, 47, 52, 54, 55] on the SiGe HBT devices under different stress conditions ($I_{\mathrm{E,stress}}$, $V_{\mathrm{CB,stress}}$) distributed uniformly across the output plane in the region defined by $BV_{\mathrm{CBO}} > V_{\mathrm{CE,stress}} > BV_{\mathrm{CEO}}$. The device degradation was measured as a change in the ratio of the pre-stress to post-stress base current leakage (extracted at low to moderate injection bias from the Gummel plots) as a function of the stress time. Each stress condition used a fresh device (i.e. the devices do not accumulate damage from or experience multiple stress conditions, and thus the damage is termed as non-accumulated), and the devices were stressed for the same specified stress-time across all stress conditions on the output plane. This non-accumulated stress degradation was measured at specified stress-time intervals and plotted (with a color scale) across the output plane to create an experimental stress-induced device degradation map or spectrum for the different stress time intervals (similar to the simulated stress damage map shown in [55]). These maps show an evolution of the non-accumulated stress-damage with stress-time across the entire output plane. Similar maps (with a color scale) for the impact-ionization coefficient (or M-1) can also be generated across the output plane for each stress condition, providing insight into how M-1 correlates to the stress damage, and thus how device design can change the stress degradation map on the output plane. Both of these methods can be used to generate contour plots using a color scale, if the experimental points for the stress-map on the output plane are more dense with finer intervals.

The second part of this study uses two different accumulated stress-test methods to characterize the nature of annealing in a SiGe HBT across the entire output plane. One of these methods of accumulated stress testing is performed at a constant stress voltage as shown in Fig. 57(a), while the stress current is swept in an increasing (or decreasing) order on the same device. During this sweep, M-1 decreases monotonically with increasing current due to Kirk effect, while the junction temperature (T_{J}) rises with increasing power dissipation

in the device, further reducing M-1. In the other method depicted in Fig. 57(b), testing is performed at a constant stress current, while the stress voltage is swept in an increasing or decreasing order on the same device. During this sweep, M-1 increases monotonically with increasing stress voltage, while T_J rises with increasing power dissipation. The opposite is true for the reverse sweeps (with decreasing stress current or voltage).

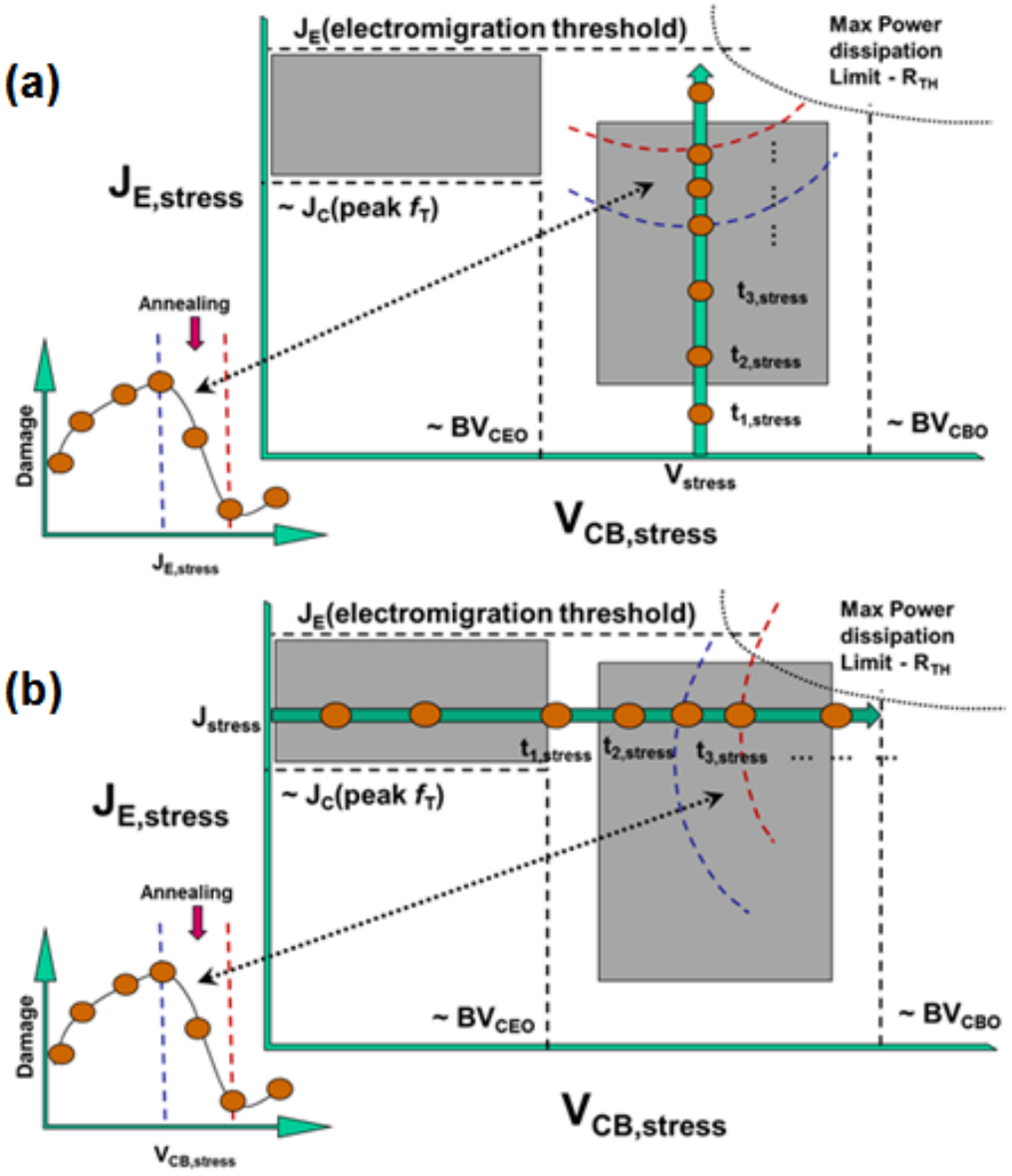

Figure 57: Accumulated stress testing methods and criterion used to identify annealing regions on the output plane. (a) Constant-voltage accumulated stress testing. (b) Constant-current accumulated stress testing. The gray colored regions indicate two commonly studied stress reliability regions at high-V (medium- to high-I) and low-V (high-I).

At each stress condition (defined by an unique stress current and voltage) during a sweep, the device undergoes stress for the same specified stress time. The damage is accumulated from the multiple stress conditions along the entire sweep. The device degradation was measured as a change in the post-stress base current leakage after each stress condition, normalized to the initial pre-stress base current (for the fresh unstressed device before the stress sweep begins). Based on the above definition of device degradation, "annealing" is identified in a region of the sweep if the damage from a specific stress condition gives lower value when compared to the damage from the previous stress condition along the sweep (i.e. net damage reduces). This concept of identifying an annealing regions on the output plane during a dynamic stress-sweep is outlined in the Figs. 57(a) and (b).

These two test accumulated stress-methods are complementary and orthogonal in their nature of studying accumulated stress damage across the output plane and providing information about the location of annealing regions. The results from the above test methods can be combined to provide evidence about the existence of annealing regions around the entire output plane, and to identify factors that determine the locations of these regions in relation to the traditional SOA of the device. The results from these test methods are analyzed and interpreted based on the R-D model proposed earlier in Chapter 2. All measurements were performed at 300 K unless otherwise mentioned.

9.3 SOA Construction Method for Comparing Device Reliability

Several inadequacies of the SOA as defined in [52] have been outlined in Chapter 2. The SOA construction to assess mixed-mode stress reliability from non-accumulated stress testing as presented in [55] assumes that the M-1, T_J for the device and the degradation criterion (ΔI_B(post-stress)/ ΔI_B(pre-stress)) are continuous functions of the stress conditions across the relevant domains on the output plane, as the rate of underlying reliability degradation mechanisms are also continuously variable functions across the same domain of stress-conditions.

Based on these valid assumptions, colored contours superimposed on the output characteristics can be constructed as an example of a 3-D (three-dimensional) contour map, and by using interpolation from these plots continuous line contours of M-1 values can be generated for the device in relevant regions of the output plane as shown in Figs. 58(a) and (b). Using this tool, M-1 for two different devices with matched performance (e.g., *npn* and *pnp* from the same C-SiGe BiCMOS process) can be compared to understand the relative strength of the impact-ionization process in these devices, how it changes across the output plane for different stress conditions (as shown in Fig. 58). Since M-1 is a measure of how strong is the impact-ionization process at a point on the output plane, it is a more fundamental measure of hot-carrier induced degradation processes operative within the device, and is also independent of device geometry (typically to first order). This information provides insight into the contributions of M-1 to the device reliability degradation by correlation to the stress damage maps. This tool can be used to study reliability implications of the device design during technology development causing changes in M-1 for the same process technology, or for devices with closely comparable performance but from different process technologies. Additionally, M-1 is a function of the device design (junction doping profiles) but is fairly independent of the device layout for most practical purposes. Thereby, it is easier to compare M-1 across different device geometries from the same or different device technologies.

As the devices undergo non-accumulated mixed-mode stress at different stress conditions on the output plane, the calculated device degradation using the base current change criterion as defined earlier can be represented with a graduated color scale to generate a stress damage map for all the stress conditions used across the output plane, as shown in Fig. 59(a). The calculated damage (or color) is a function of the stress condition and time, device geometry, temperature, and device design (through M-1, R_{TH}, T_{J}, etc.), thereby making generalized interpretations and comparisons very difficult across large regions of the output plane. However, this can be simplified by constructing few graded line contours

from interpolation of the stress map by connecting nearby stress points with equal device degradation, as shown in Fig. 59(a) superimposed with the corresponding stress map. The damage contours can be interpreted as operating conditions with equal reliability or device degradation for the specific accelerated stress time and method used. This provides a procedure to estimate device reliability for a specific stress condition, as well as to choose an operating condition for a specific desired reliability. Taken together, this provides an easier method to construct the practical SOA of the device across different regions of the output plane with a certain desirable reliability. This contour method of the SOA construction can also be used to compare reliability of multiple devices with comparable performance when superimposed upon each other (as shown in Fig. 59(b)). In general, this method can be extended to be used in the comparison of superimposed SoA for different device geometries, across different temperatures for the same device, to show stress-time evolution of reliability, or across different device technologies with comparable performance. This proposed method is a much simpler tool to be used for estimation and comparison of device reliability, as opposed to a full-blown detailed non-accumulated stress map with a dense array of stress conditions which would require a significantly larger set of measurements. The detailed stress map includes random and process technology variations (each point being measured on a different device). Thus, the damage contour method of comparison allows a way of averaging over the random device to device variations between any two devices used in the measurements, as it interpolates the damage between these two stress conditions, assuming the calculated damage to be a continuous function on the output plane.

As an example, in Fig. 58, although the *npn* device has a higher M-1 at lower voltages which explains a lower BV_{CEO} for the device, however the *pnp* has a higher M-1 at higher voltages, indicating that the M-1 for the *pnp* increases much more sharply with increasing voltage and so has a higher doping compared to the *npn* at the collector-base junction. The *pnp* has higher doping and impact-ionization coefficient M-1, thereby explaining a higher hot-carrier induced device damage compared to the *npn* device.

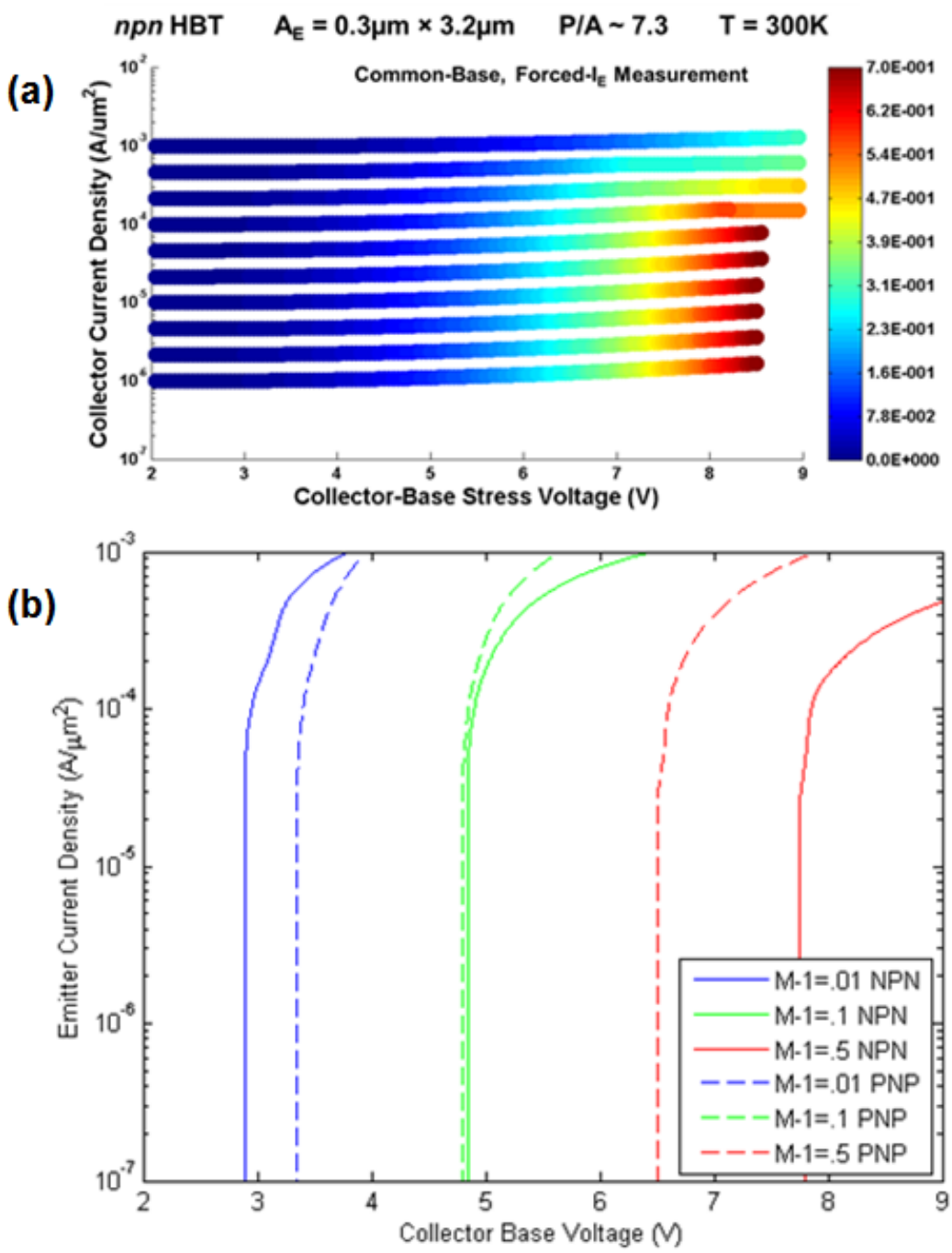

Figure 58: (a) Plot of *M*-1 superimposed with a color scale on the output characteristics of a device, resulting in a color map or contours. (b) Plots of continuous *M*-1 line contours with graduent calculated from the same measurement, compared for an *npn* and *pnp* device with comparable performance from the same C-SiGe process technology.

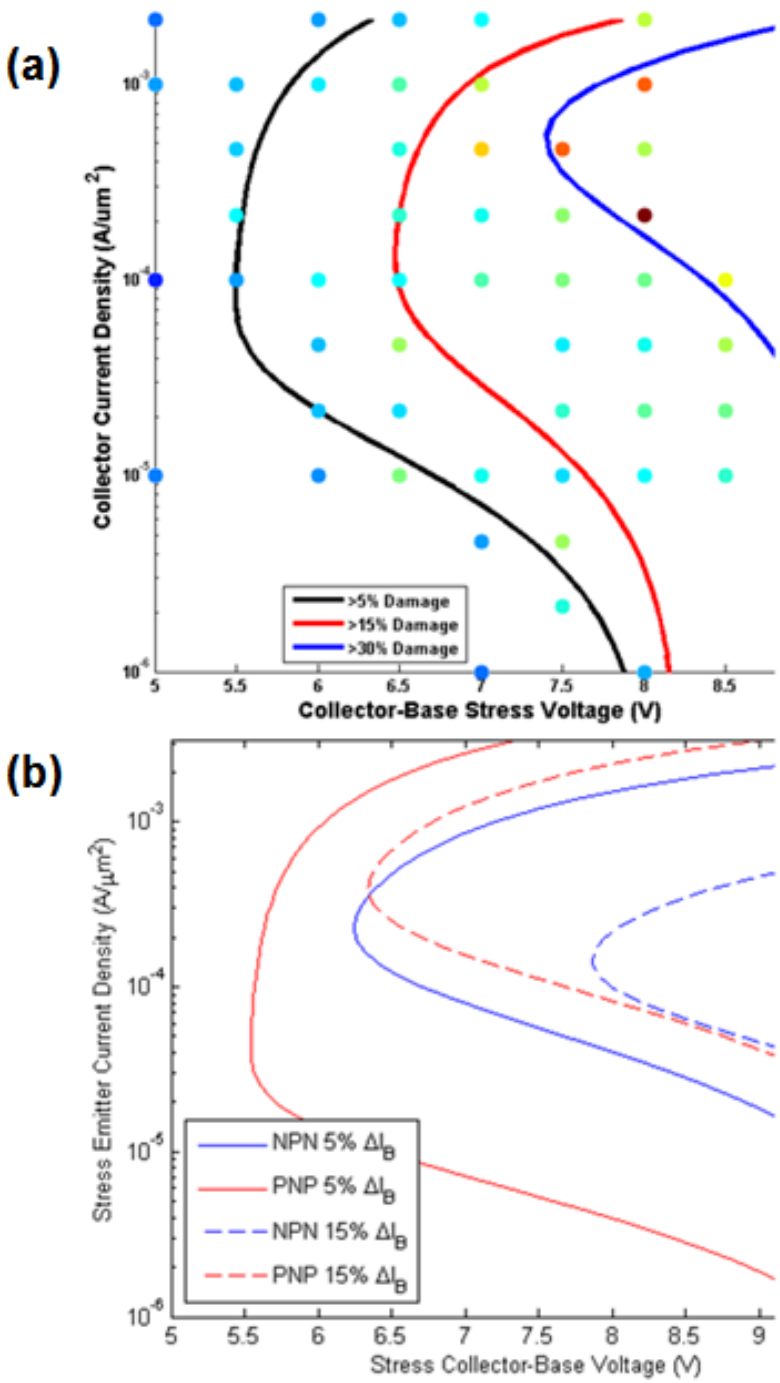

Figure 59: (a) Mixed-mode stress damage spectrum with a color scale, superimposed with the calculated contours for different amounts of device degradation, using the base current leakage criterion. (b) Comparison of the degradation contours for two different devices with different levels of device degradation using the same criterion, stress time, ambient temperature, and device geometry. The results are for an *npn* and *pnp* device with comparable performance from the same C-SiGe BiCMOS technology.

9.4 Results and Analysis of Annealing Studies

For the purpose of this study, the following discussion will be primarily focused on the degradation quantitatively derived from the base leakage in the forward gummel measurements. Mostly, the damage and annealing regions will be identified based on how the forward gummel base-current evolves with the stress-condition and time. However, correlation with the inverse gummel behavior will be introduced as needed to aid the analysis.

The reaction-diffusion model for hot-carrier induced damage creation (as outlined in the equation below) during an accelerated reliability test (like mixed-mode stress) can be broadly defined (simplistically) as a reaction between the generated hot-carrier from the accelerated stress and a Si-H bond at a oxide-silicon interface, leading to creation of a Si dangling bond, along with the diffusion of free H atoms (or species) away from the reaction interface. The reverse reaction can be outlined as the damage annealing process, which will depend on the presence of a Si dangling bond (or a free trap level) and the diffusion of H atoms (or species) to the location of the trap interface, and recombining to form a S-H bond.

$$Hot - Carrier(HC) + Si - H(bond) \rightleftharpoons Si - (danglingbond) + H(free) \tag{2}$$

The overall reaction rate (based on forward and reverse reaction rate constants) as given below is a function of the individual concentrations of the reactants and the products at any specific instance of stress-time.

$$k = \frac{K_R}{K_F} = \frac{[HC][Si - H]}{[Si-][H]} \tag{3}$$

If the forward reaction is dominant over the reverse reaction, net damage creation happens in the form of dangling bonds or traps as a function of stress-time. If the reverse reaction is dominant, annealing of the traps is observed with increasing stress-time. At any instance of time, both the forward and reverse reactions are operative concurrently with rates dependent on the device structure, processing steps, device temperature, along

with concentration of the reactants and products in the reaction model above. One way device temperature directly plays a dominant role is through the diffusion and transport of H atoms towards and away from the damage interface, thus controlling the reverse reaction or annealing rate of the damage. In any generalized operating trajectory of the device on the output plane, either one of the forward or reverse reactions will dominate based on the hot-carrier concentration at the oxide-silicon interface, concentration of existing dangling bonds or interface traps, and the supply of H atoms for annealing of the traps. The dominant reaction will lead to a manifestation of net annealing or damage at the oxide-silicon interfaces as calculated from the forward or inverse gummel plots. A dominant reverse reaction or resultant annealing requires high concentration of traps and available H atoms. It is already known that the forward gummel manifests damage/annealing at the emitter-base spacer oxide-silicon interface, while the reverse gummel is an indicator for the shallow-trench oxide-silicon interface [52]. The results shown in the following part is representative of devices in the technology under consideration, and has been verified on atleast three or more devices for each plot shown. The analysis presented uses trends and correlation of trends from the accumulated stress sweeps, instead of specific numbers which can be very sensitive to device-to-device variations. Each accumulated stress-sweep was performed on a fresh device.

As shown in Fig. 60 for the 200 GHz *npn* SiGe HBT, three different devices are subjected to constant-voltage accumulated stress sweep as defined in Section 9.2. It is important to be reminded here that while M-1 monotonically decreases along the sweep, the device temperature increases with higher power dissipation. The stress times the devices saw at each mixed-mode stress point are different (100, 500, and 1000 seconds respectively). This would lead to different amount of damage creation (or interface trap concentration) within the devices. However, the location of the observed annealing region boundaries still remains exactly same for all the three sweeps. This indicates to one single factor which is common among the three devices at a common stress condition during the sweeps, which is

the device or junction temperature. It is the device temperature that drives the diffusion of H atom (or species) towards the interface, leading to strong reverse reaction and annealing being dominant. It is safe to assume that these devices will show significant self-heating for the stress conditions required to observe the onset of annealing. When the inverse gummel degradation for the same devices (from Fig. 60) is plotted in Fig. 61, a very close correlation in the location (stress current) of the annealing region boundaries for all three sweeps is observed, thereby confirming the assertions from FG results. According to the R-D model presented earlier, both high concentration of traps and H atoms are needed for annealing. In this case it is clear that the excess concentration of H species resulting from the diffusion process drives the reverse reaction to be dominant.

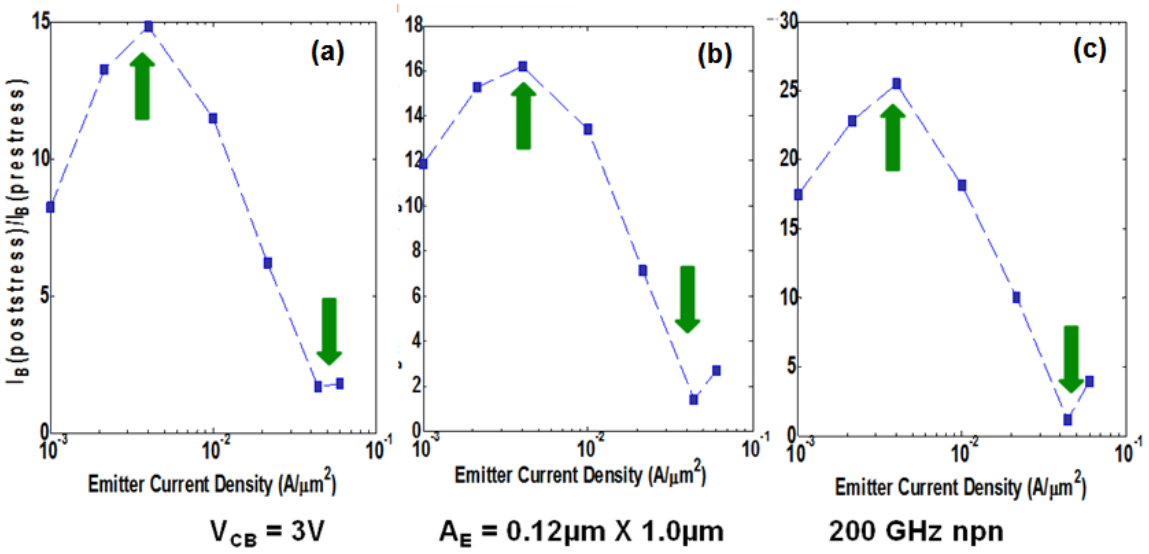

Figure 60: The plots show constant-voltage stress sweep damage (as calculated from the forward gummel of these devices) and annealing on a 200 GHz *npn* device when undergoing accumulated mixed-mode stress for (a) 100, (b) 500, and (c) 1000 seconds at each stress condition. The location of the annealing regions are indicated in each plot (measured on a different device), indicating the location of the annealing regions on the output plane matches exactly for each of these devices (as measured at the emitter-base spacer oxide-silicon interface using forward gummel).

As shown in Fig. 62, two devices were subjected to the constant-current accumulated stress sweep as defined in Section 9.3. It is important to be reminded that while *M*-1 or hot-carrier density monotonically increases along the sweep, the device temperature also

increases with higher power dissipation. The stress times the devices saw at each mixed-mode stress points are 100 and 1000 seconds respectively. This would lead to different amount of damage creation (or interface trap concentration) within the devices. The first important observation is the fact that both of these plots with same stress conditions but different stress times show the existence of two distinct annealing regimes during each sweep. No such concept about the existence of two annealing regions at a specific current or voltage have been observed for the same devices earlier [52].

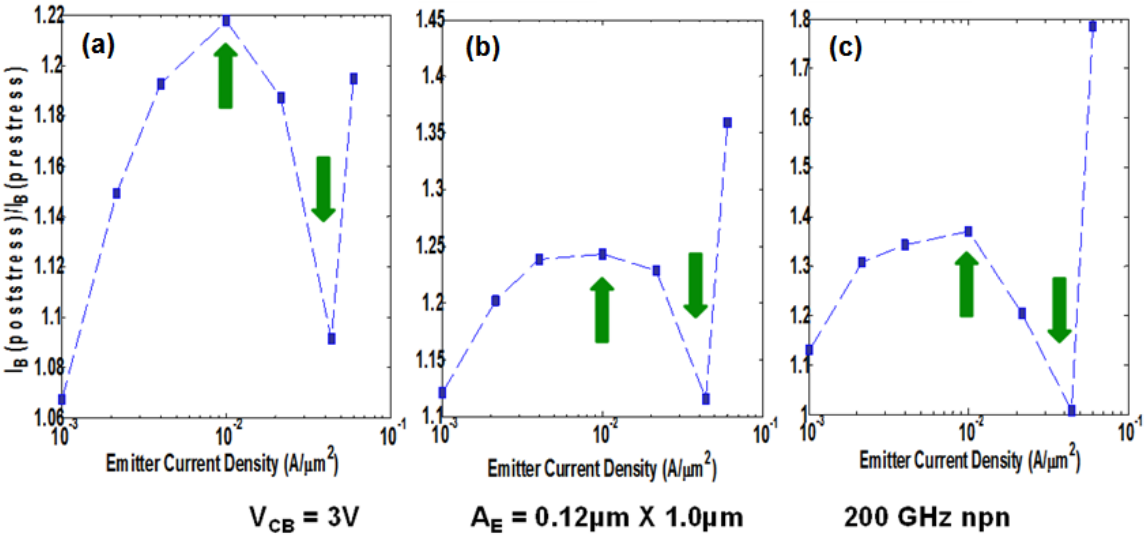

Figure 61: The plots show constant-voltage stress sweep damage (as calculated from the inverse gummel of the devices in Fig. 60) and annealing on a 200 GHz *npn* device when undergoing accumulated mixed-mode stress for (a) 100, (b) 500, and (c) 1000 seconds at each stress condition. The location of the annealing regions are outlined in each plot (measured on a different device), indicating the location of the annealing regions on the output plane matches up exactly for each of these devices (as measured at the shallow-trench oxide-silicon interface using reverse gummel).

The onset of the first annealing region (indicated by a red arrow) for the 100 seconds stress sweep in Fig. 62(a) moves to a lower stress voltage with the 1000 seconds stress sweep as shown in Fig. 62(b). However, the onset of the second annealing region (indicated by a green arrow) in Fig. 62(a) moves to a higher stress voltage with the 1000 seconds stress sweep as shown in Fig. 62(b). Since the device has a higher concentration of traps in Fig. 62(b) at the onset of the first annealing region, it can be safely said that the dominant reverse

reaction rate in the first annealing region is driven by an excess interface trap concentration. Once enough traps have been annealed during the stress sweep, and there are much less number of traps available to be annealed, the reverse reaction rate slows down, leading to again a dominant forward reaction rate or trap creation process.

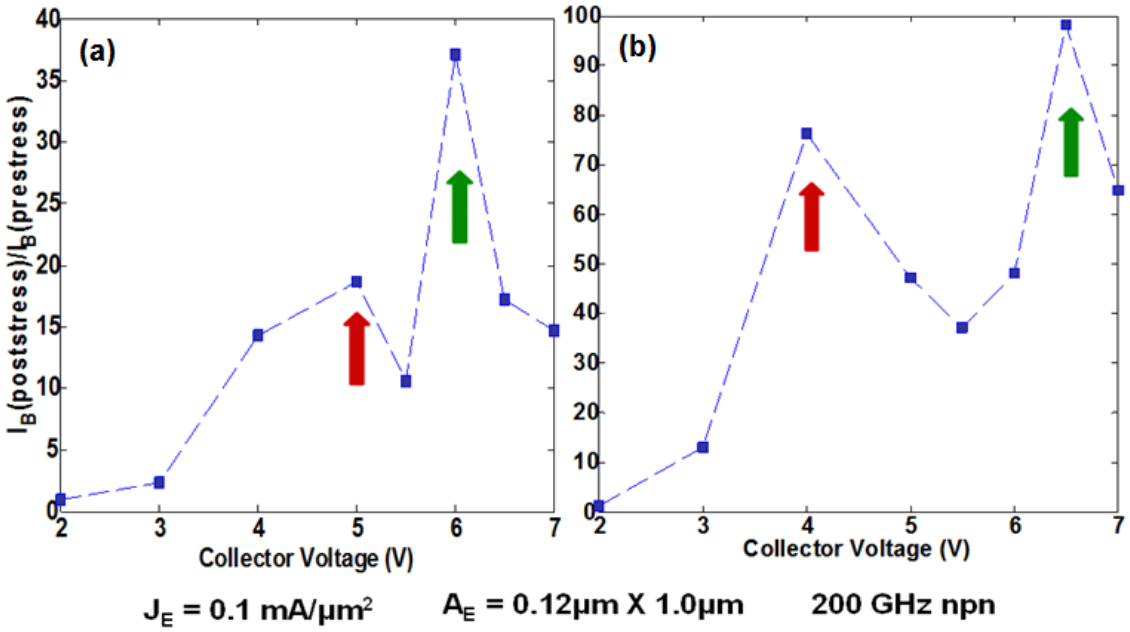

Figure 62: The plots show constant-current stress sweep damage (as calculated from the forward gummel of the devices) and annealing on a 200 GHz *npn* device when undergoing accumulated mixed-mode stress for (a) 100, and (b) 1000 seconds at each stress condition. The location of the annealing regions are outlined in each plot (measured on a different device), indicating the presence of two annealing regions. The location of onset of the two annealing regions on the output plane moves in opposite directions between these two sweeps, indicating the presence of two different dominant factors responsible for annealing in each region of the device.

As the stress sweep moves to a much higher voltage along the sweep, the device heats up enough with increased power dissipation to create sufficient H atoms to diffuse to the interface for a dominant reverse reaction, and thus a second annealing region is observed due to excess H diffusion. Since the 1000 seconds stress creates more damage than the 100 seconds stress, it requires a higher temperature within the device for a high enough diffusion rate of H atoms reaching the interface to create a dominant reverse reaction rate for observing annealing. Thus the onset of the second annealing region happens at a higher

voltage (or device temperature) for the 1000 second stress sweep compared to the 100 seconds sweep.

In Fig. 63, the constant-current sweep is shown at a much higher current for which there is significant self-heating even at lower stress voltages. The hot-carrier generation is from both impact-ionization and auger processes. The onset of the first annealing regime is at the same voltage for both 100 and 1000 seconds sweeps, indicating a dependence on device temperature (which controls the H diffusion rate), once there is enough trap concentration created within the devices during the sweep. The onset of the second annealing region (green arrow) moves to a lower stress voltage for the 1000 seconds sweep, as a much higher concentration of traps (which controls the reverse reaction) are created earlier during the sweep. This is a proof of the fact that the second annealing region in this case is driven by a critical concentration of traps, possibly because there is already high enough availability of diffusing H atoms resulting from the high device temperature during the stress condition.

Similar results and trends were observed on lower performance devices ($f_T < 50$ GHz), lending further validity to the presented analysis. However, the overall damage and annealing changes observed for such lower performance devices are much less due to a significantly weaker impact-ionization process and lower hot-carrier generation rates, as well as longer distances the hot-carriers have to propagate to create damage. Overall, these case studies prove that a dominant net reverse reaction rate (manifested as annealing of hot-carrier damage) can critically depend on two different important factors as proposed by the reaction-diffusion criteria for hot-carrier induced damage modeling presented above. First being the availability of a critical concentration of traps at the beginning of a stress condition, when there are enough H atoms (or species like H_2). Second being the availability of a critical number of H atoms diffusing to the interface traps (diffusion rate is primarily dependent on the device temperature), when there is already enough trap concentration available at the start of a stress test. Both of these critical cases of reaction kinetics has been presented and analyzed in details through Figs. 60, 61, 62, 63.

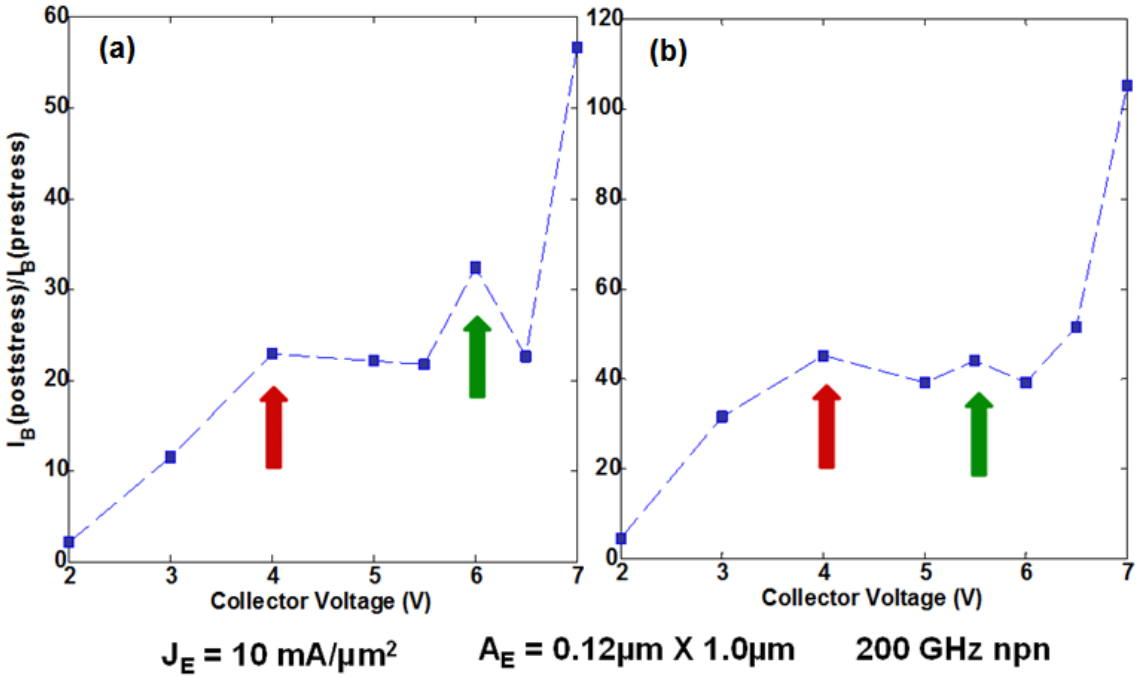

Figure 63: The plots show constant-current stress sweep damage (as calculated from the forward gummel of the devices) and annealing on a 200 GHz *npn* device when undergoing accumulated mixed-mode stress for (a) 100, and (b) 1000 seconds at each stress condition. The location of the annealing regions are outlined in each plot (measured on a different device), indicating the presence of two annealing regions. The location of the two annealing regions on the output plane changes in opposite directions between these two sweeps indicating presence of two different factors responsible for annealing.

Taken together, this study highlights the most important factors for predictive TCAD modeling of hot-carrier induced damage and annealing using the R-D formalism. It has been highlighted earlier that the forward reaction rate can be modeled well in TCAD using proper models and careful calibration of hot-carrier rates and interface physics [55]. However, for the satisfactory implementation of a net reverse reaction rate or annealing, accurate modeling of both the rate of hot-carrier generated trap concentration at the interface, as well as the H atom concentrations are required. The H atom concentration driving the annealing process is dependent on the H diffusion rate which is a strong function of temperature. Hence, an empirical model for the time dependent H concentration (for a given process technology and temperature) to fit the annealing data, or a functional model for the H diffusion near the interface traps would be required for fully predictive TCAD

implementation of the R-D model for hot-carrier induced damage creation and annealing.

9.5 Summary

This study first highlights the inadequacies in the traditional SOA construction of SiGe HBTs from traditional *dc* stress methods, and important factors that needs to be taken into account (like device geometry, temperature, accelerated stress method, degradation criterion, etc.) while defining the SOA. New contour methods are proposed to compare impact-ionization and hot-carrier degradation across the output plane between two devices from the same technology but with different geometry, or devices with comparable performances from different technologies. These tools can be used for device design and reliability optimization during technology development. Circuit designers can use these tools to implement device geometries with better reliability for a specific topology and target application, based on the bias trajectory seen during device operation on the output plane.

In the second part of the study, constant-voltage and constant-current accumulated stress sweeps are used to study validity of the R-D model for hot-carrier generation and particularly annealing of the hot-carrier induced damage. The two critical factors determining the annealing (or reverse reaction) rate are validated through the case studies. Overall, this study provides the first experimental proof reported for the reaction-diffusion model for annealing in SiGe HBTs, and the critical factors controlling the annealing reaction rates. This is of significant importance towards the success of TCAD modeling for hot-carrier reliability on the output plane, particularly in the ongoing effort of using the R-D model as introduced earlier in Chapter 2.

All these methods can also be combined to develop and calibrate a reliability compact model, to be further integrated into a circuit design environment, and thus can potentially lend more flexibility to circuit designers.

CHAPTER 10

CONCLUSIONS AND FUTURE WORK

The objective of this research work is to investigate and gain new understanding on how device design couples with both performance scaling and reliability for mixed-signal applications (RF and analog), and using this knowledge to enhance predictive modeling of performance and reliability for these devices. These results are tied together with a guiding theme: to develop a holistic understanding about how the device design factors influence both performance scaling and reliability. In addition, this work provides methods for using that knowledge to enable predictive modeling of performance and reliability for mixed-signal devices in BiCMOS technologies. The preliminary research remains in the device realm and fill some of the key existing gaps in the understanding of predictive-modeling for performance scaling and reliability. These results can be further leveraged by engineers working on research and development of new device technologies to predictively model both device performance and reliability in an integrated TCAD framework. In this chapter, the contributions of this dissertation are summarized, and suggestions for future research are provided.

10.1 Contributions

The contributions of this work can be summarized as follows:

1. Development of an improved optimization strategy and integrated TCAD framework for predictive modeling of the performance scaling complementary SiGe HBTs. This includes the first feasibility study of a 200 GHz SiGe HBT and demonstration of a complementary SiGe HBT Roadmap using TCAD. A method of optimizing complementary devices for performance matching over a range of bias conditions is also demonstrated.

2. Identifying factors influencing the predictive nature of the simulated FoM (like output

conductance) of SiGe HBTs, and verification through experimentation.

3. Predictive-modeling of the reverse-biased emitter-base junction tunneling current mechanisms in SiGe HBTs and how that couples to performance scaling. Reliability degradation due to tunneling stress and its dependence on performance scaling is also introduced.

4. Providing insight into the electrothermal constraints arising out of scaling vertical SiGe HBTs on thick-film SOI, and how that influences the *dc* and RF SOA of the device. Conventional and newly proposed measurement techniques are used to study several *dc* and RF FoM for the devices.

5. Provides the first experimental proof for the physical mechanisms involved in the mixed-mode stress damage of *pnp* SiGe HBTs.

6. Demostration of the best cryogenic performance reported to date of a "best-of-breed" *npn* SiGe HBT, and its comparison with the cryogenic performance of a *pnp* from the "best-of-breed" complementary SiGe HBT technology.

7. Proposes new methods to compare impact-ionization coefficient and hot-carrier reliability of devices more generally across device technologies or geometries in an integrated way. In addition, it provides experimental validation of the R-D model for the mechanisms driving the annealing of hot-carrier damage.

10.2 Future Work

The research work presented in this thesis establishes several interesting opportunities for future work:

1. Developing experimental comparisons between RF and *dc* SOA to enable predictive TCAD modeling of device reliability.

2. Studying high-temperature performance of "best-of-breed" *npn* SiGe HBT devices and comparison with "best-of-breed" complementary SiGe HBTs.

3. Developing more predictive reliability compact models using experiments and TCAD modeling.

4. Implementation of an annealing model for estimating hot-carrier damage within a TCAD framework.

5. TCAD modeling and device optimization of SiGe HBTs for power-amplifier applications.

REFERENCES

[1] P. S. Chakraborty, K. A. Moen, and J. D. Cressler, "An Investigation on the Optimization and Scaling of Complementary SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 60, pp. 34–41, Jan 2013.

[2] P. S. Chakraborty, A. S. Cardoso, B. R. Wier, A. P. Omprakash, J. D. Cressler, M. Kaynak, and B. Tillack, "A 0.8 THz f_{MAX} SiGe HBT Operating at 4.3 K," *IEEE Electron Device Letters*, vol. 35, pp. 151–153, Feb 2014.

[3] P. S. Chakraborty, S. J. Horst, K. A. Moen, M. Bellini, and J. D. Cressler, "An investigation of electro-thermal instabilities in 150 GHz SiGe HBTs fabricated on SOI," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 141–144, Oct 2010.

[4] P. S. Chakraborty, K. A. Moen, J. D. Cressler, H. Ho, H. Yasuda, B. Eklund, and R. Wise, "Predictive Output Conductance Modeling of SiGe HBTs Using TCAD," in *Proceedings of the International Silicon-Germanium Technology and Devices Meeting*, pp. 1–2, May 2010.

[5] P. S. Chakraborty, K. A. Moen, and J. D. Cressler, "Predictive TCAD modeling of the scaling-induced, reverse-biased, emitter-base tunneling current in SiGe HBTs," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 13–16, Oct 2011.

[6] P. S. Chakraborty, A. C. Appaswamy, P. K. Saha, N. K. Jha, J. D. Cressler, H. Yasuda, B. Eklund, and R. Wise, "Mixed-Mode Stress Degradation Mechanisms in *pnp* SiGe HBTs," in *Proceedings of the IEEE International Reliability Physics Symposium (IRPS)*, pp. 83–88, April 2009.

[7] P. S. Chakraborty, K. Moen, M. Bellini, and J. D. Cressler, "Investigation of the device design challenges and optimization issues associated with complementary SiGe HBT scaling," in *Proceedings of the International Semiconductor Device Research Symposium*, pp. 1–2, Dec 2009.

[8] J. D. Cressler, ed., *Silicon Heterostructure Handbook: Materials, Fabrication, Devices, Circuits, and Applications of SiGe and Si Strained-Layer Epitaxy*. Boca Raton, FL: CRC Taylor & Francis, 2006.

[9] J. D. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistors*. Boston: Artech House, 2003.

[10] T. H. Ning, "Why BiCMOS and SOI BiCMOS?," *IBM Journal of Research and Development*, vol. 46, pp. 181–186, March 2002.

[11] The International Technology Roadmap for Semiconductors, *Radio Frequency and Analog Mixed-Signal Technologies*, 2013.

[12] S. K. Wiedmann, "Potential of Bipolar Complementary Device/Circuit technology," in *Technical Digest of the IEEE International Electron Devices Meeting*, vol. 33, pp. 96–99, 1987.

[13] J. Warnock, J. D. Cressler, J. Burghartz, D. Harame, K. Jenkins, and C.-T. Chuang, "High Performance Complementary Bipolar Technology," in *Digest of Technical Papers for the IEEE Symposium on VLSI Technology*, pp. 75–76, May 1993.

[14] D. M. Monticelli, "The Future of Complementary Bipolar," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 21–25, Sept 2004.

[15] J. D. Cressler, "Issues and Opportunities for Complementary SiGe HBT Technology," in *Proceedings of the ECS-SiGe Meeting*, pp. 893–911, 2006.

[16] E. Zhao, A. K. Sutton, B. M. Haugerud, J. D. Cressler, P. W. Marshall, R. A. Reed, B. El-Kareh, S. Balster, and H. Yasuda, "The effects of radiation on 1/f noise in complementary (npn+pnp) SiGe HBTs," *IEEE Transactions on Nuclear Science*, vol. 51, pp. 3243–3249, Dec 2004.

[17] M. Bellini, B. Jun, A. K. Sutton, A. C. Appaswamy, P. Cheng, J. D. Cressler, P. W. Marshall, R. D. Schrimpf, D. M. Fleetwood, B. El-Kareh, S. Balster, P. Steinmann, and H. Yasuda, "The Effects of Proton and X-Ray Irradiation on the DC and AC Performance of Complementary (npn + pnp) SiGe HBTs on Thick-Film SOI," *IEEE Transactions on Nuclear Science*, vol. 54, pp. 2245–2250, Dec 2007.

[18] B. El-Kareh, S. Balster, W. Leitz, P. Steinmann, H. Yasuda, M. Corsi, K. Dawoodi, C. Dirnecker, P. Foglietti, A. Haeusler, P. Menz, M. Ramin, T. Scharnagl, M. Schiekofer, M. Schober, U. Schulz, L. Swanson, D. Tatman, M. Waitschull, J. W. Weijtmans, and C. Willis, "A 5V complementary-SiGe BiCMOS technology for high-speed precision analog circuits," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 211–214, Sept 2003.

[19] B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, O. Fursenko, M. Glante, K. Glowatzki, A. Gregor, U. Haak, W. Hoppner, D. Knoll, R. Kurps, S. Marschmeyer, S. Orlowski, H. Rucker, P. Schley, D. Schmidt, R. Scholz, W. Winkler, and Y. Yamamoto, "A complementary BiCMOS technology with high speed npn and pnp SiGe:C HBTs," in *Technical Digest of IEEE International Electron Devices Meeting*, pp. 5.2.1–5.2.4, Dec 2003.

[20] J. Duvernay, G. Borot, P. Chevalier, D. Dutartre, R. Pantel, L. Rubaldo, T. Schwartzmann, B. Vandelle, and A. Chantre, "An experimental and simulation study of pnp Si/SiGeC HBTs using box-like Ge profiles," in *Proceedings of the European Solid State Device Research Conference*, pp. 346–349, Sept 2007.

[21] D. Knoll, B. Heinemann, Y. Yamamoto, H.-E. Wulf, and D. Schmidt, "PNP SiGe: C HBT Optimization in a Low-Cost CBiCMOS Process," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 30–33, Sept 2007.

[22] J. Duvernay, F. Brossard, G. Borot, L. Boissonnet, B. Vandelle, L. Rubaldo, F. Deleglise, G. Avenier, P. Chevalier, B. Rauber, D. Dutartre, and A. Chantre, "Development of a self-aligned pnp HBT for a complementary thin-SOI SiGeC BiCMOS technology," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 34–37, Sept 2007.

[23] T. Tominari, M. Miura, H. Shimamoto, M. Arai, Y. Yoshida, H. Sato, T. Aoki, H. Nonami, S. Wada, H. Hosoe, K. Washio, and T. Hashimoto, "A 10V complementary SiGe BiCMOS foundry process for high-speed and high-voltage analog applications," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 38–41, Sept 2007.

[24] E. J. Preisler, L. Lao, J. Zheng, P. Hurwitz, and M. Racanelli, "Low cost, highly flexible complementary bipolar transistors compatible with 0.18 or 0.13 μm CMOS technology," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 101–104, Oct 2008.

[25] J. A. Babcock, G. Cestra, W. van Noort, P. Allard, S. Ruby, J. Tao, R. Malone, A. Buchholz, N. Lavrovskaya, W. Yindeepol, C. Printy, J. Ramdani, A. Labonte, H. McCulloh, Y. Leng, P. McCarthy, D. Getchell, A. Sehgal, T. Krakowski, S. Desai, C. Joyce, P. Hojabri, and S. Decoutere, "CBC8: A 0.25 μm SiGe-CBiCMOS technology platform on thick-film SOI for high-performance analog and RF IC design," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 41–44, Oct 2010.

[26] G. Zhang, J. D. Cressler, G. Niu, and A. Pinto, "A comparison of npn and pnp profile design tradeoffs for complementary SiGe HBT Technology," *Solid-State Electronics*, vol. 44, no. 11, pp. 1949 – 1954, 2000.

[27] T. Chen, J. Babcock, Y. Nguyen, W. Greig, N. Lavrovskaya, T. Thibeault, S. Ruby, S. Adler, T. Krakowski, J. Kim, and A. Sadovnikov, "Footprint design optimization in SiGe BiCMOS SOI technology," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 208–211, Oct 2008.

[28] B. G. Malm, E. Haralson, T. Johansson, and M. Ostling, "Self-heating effects in a BiCMOS on SOI technology for RFIC applications," *IEEE Transactions on Electron Devices*, vol. 52, pp. 1423–1428, July 2005.

[29] J.-S. Rieh, D. Greenberg, B. Jagannathan, G. Freeman, and S. Subbanna, "Measurement and modeling of thermal resistance of high speed SiGe heterojunction bipolar transistors," in *Proceedings of the IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, pp. 110–113, Sept 2001.

[30] T. Vanhoucke, H. M. J. Boots, and W. D. Van Noort, "Revised method for extraction of the thermal resistance applied to bulk and SOI SiGe HBTs," *IEEE Electron Device Letters*, vol. 25, pp. 150–152, March 2004.

[31] J. Cai, A. Ajmera, C. Ouyang, P. Oldiges, M. Steigerwalt, K. Stein, K. Jenkins, G. Shahidi, and T. Ning, "Fully-depleted-collector polysilicon-emitter SiGe-base vertical bipolar transistor on SOI," in *Digest of Technical Papers for the IEEE Symposium on VLSI Technology*, pp. 172–173, June 2002.

[32] G. Avenier, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, D. Dutartre, L. Boissonnet, F. Saguin, R. Pantel, S. Fregonese, C. Maneux, T. Zimmer, and A. Chantre, "A self-aligned vertical HBT for thin SOI SiGeC BiCMOS," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 128–131, Oct 2005.

[33] F. Sato, T. Hashimoto, H. Tezuka, M. Soda, T. Suzaki, T. Tatsumi, and T. Tashiro, "A 60-GHz f_T super self-aligned selectively grown SiGe-base (SSSB) bipolar transistor with trench isolation fabricated on SOI substrate and its application to 20-Gb/s optical transmitter ICs," *IEEE Transactions on Electron Devices*, vol. 46, pp. 1332–1338, Jul 1999.

[34] K. Washio, E. Ohue, H. Shimamoto, K. Oda, R. Hayami, Y. Kiyota, M. Tanabe, M. Kondo, T. Hashimoto, and T. Harada, "A 0.2- μm 180-GHz-f_{max} 6.7-ps-ECL SOI/HRS self-aligned SEG SiGe HBT/CMOS technology for microwave and high-speed digital applications," *IEEE Transactions on Electron Devices*, vol. 49, pp. 271–278, Feb 2002.

[35] I. Z. Mitrovic, O. Buiu, S. Hall, D. M. Bagnall, and P. Ashburn, "Review of SiGe HBTs on SOI," *Solid-State Electronics*, vol. 49, no. 9, pp. 1556 – 1567, 2005.

[36] W. Liu and A. Khatibzadeh, "The collapse of current gain in multi-finger heterojunction bipolar transistors: its substrate temperature dependence, instability criteria, and modeling," *IEEE Transactions on Electron Devices*, vol. 41, pp. 1698–1707, Oct 1994.

[37] T. Vanhoucke and G. A. M. Hurkx, "Unified electro-thermal stability criterion for bipolar transistors," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 37–40, Oct 2005.

[38] W. D. Van Noort and R. Dekker, "Thermal resistance of HBTs on bulk Si, SOI and glass," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 129–132, Sept 2003.

[39] N. Nenadovic, V. d'Alessandro, L. K. Nanver, N. Rinaldi, H. Schellevis, and J. W. Slotboom, "Analytical formulation and electrical measurements of self-heating in silicon BJTs," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 24–27, Sep-Oct 2002.

[40] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Transactions on Electron Devices*, vol. 39, pp. 331–338, Feb 1992.

[41] A. Schenk, "Rigorous theory and simplified model of the band-to-band tunneling in silicon," *Solid-State Electronics*, vol. 36, no. 1, pp. 19 – 34, 1993.

[42] D. Lagarde, P. Chevalier, T. Schwartzmann, and A. Chantre, "Band-to-band tunneling in vertically scaled SiGe:C HBTs," *IEEE Electron Device Letters*, vol. 27, pp. 275–277, April 2006.

[43] J. A. del Alamo and R. M. Swanson, "Forward-bias tunneling: A limitation to bipolar device scaling," *IEEE Electron Device Letters*, vol. 7, pp. 629–631, Nov 1986.

[44] A. Neviani, P. Pavan, A. Nardi, A. Chantre, L. Vendrame, and E. Zanoni, "Hot-carrier degradation and oxide charge build-up in self-aligned etched-polysilicon npn bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 44, pp. 2059–2063, Nov 1997.

[45] J. J. Liou, "Modeling the tunnelling current in reverse-biased p/n junctions," *Solid-State Electronics*, vol. 33, no. 7, pp. 971 – 972, 1990.

[46] G. Zhang, J. D. Cressler, G. Niu, and A. J. Joseph, "A New "Mixed-Mode" Reliability Degradation Mechanism in Advanced Si and SiGe Bipolar Transistors," *IEEE Transactions on Electron Devices*, vol. 49, pp. 2151–2156, Dec 2002.

[47] C. Zhu, Q. Liang, R. A. Al-Huq, J. D. Cressler, Y. Lu, T. Chen, A. J. Joseph, and G. Niu, "Damage Mechanisms in Impact-Ionization-Induced Mixed-Mode Reliability Degradation of SiGe HBTs," *IEEE Transactions on Device and Materials Reliability*, vol. 5, pp. 142–149, March 2005.

[48] T. Vanhoucke, G. A. M. Hurkx, D. Panko, R. Campos, A. Piontek, P. Palestri, and L. Selmi, "Physical Description of the Mixed-Mode Degradation Mechanism for High Performance Bipolar Transistors," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 1–4, Oct 2006.

[49] R. Krithivasan, Y. Lu, J. D. Cressler, J.-S. Rieh, M. H. Khater, D. Ahlgren, and G. Freeman, "Half-Terahertz Operation of SiGe HBTs," *IEEE Electron Device Letters*, vol. 27, pp. 567–569, July 2006.

[50] J. Yuan, J. D. Cressler, R. Krithivasan, T. Thrivikraman, M. H. Khater, D. C. Ahlgren, A. J. Joseph, and J.-S. Rieh, "On the Performance Limits of Cryogenically Operated SiGe HBTs and Its Relation to Scaling for Terahertz Speeds," *IEEE Transactions on Electron Devices*, vol. 56, pp. 1007–1019, May 2009.

[51] M. A. Oakley, B. Wier, U. S. Raghunathan, P. S. Chakraborty, and J. D. Cressler, "On the reliability of SiGe HBT cascode driver amplifiers," in *Proceedings of the IEEE Radio Frequency Integrated Circuits Symposium*, pp. 445–448, June 2014.

[52] P. Cheng, C. Zhu, A. Appaswamy, and J. D. Cressler, "A New Current-Sweep Method for Assessing the Mixed-Mode Damage Spectrum of SiGe HBTs," *IEEE Transactions on Device and Materials Reliability*, vol. 7, pp. 479–487, Sept 2007.

[53] P. Cheng, C. M. Grens, and J. D. Cressler, "Reliability of SiGe HBTs for Power Amplifiers–Part II: Underlying Physics and Damage Modeling," *IEEE Transactions on Device and Materials Reliability*, vol. 9, pp. 440–448, Sept 2009.

[54] K. A. Moen, P. S. Chakraborty, U. S. Raghunathan, J. D. Cressler, and H. Yasuda, "Predictive Physics-Based TCAD Modeling of the Mixed-Mode Degradation Mechanism in SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 59, pp. 2895–2901, Nov 2012.

[55] U. S. Raghunathan, P. S. Chakraborty, B. Wier, J. D. Cressler, H. Yasuda, and P. Menz, "TCAD Modeling of Accumulated Damage During Time-Dependent Mixed-Mode Stress," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 179–182, Sept 2013.

[56] Synopsys, Inc., Mountain View, CA, *Sentaurus Workbench*, Sep 2008.

[57] E. F. Crabbe, J. M. C. Stork, G. Baccarani, M. V. Fischetti, and S. E. Laux, "The impact of non-equilibrium transport on breakdown and transit time in bipolar transistors," in *Technical Digest of the IEEE International Electron Devices Meeting*, pp. 463–466, Dec 1990.

[58] T. Grasser, T.-W. Tang, H. Kosina, and S. Selberherr, "A review of hydrodynamic and energy-transport models for semiconductor device simulation," *Proceedings of the IEEE*, vol. 91, pp. 251–274, Feb 2003.

[59] J. Yuan and J. D. Cressler, "Design and Optimization of Superjunction Collectors for Use in High-Speed SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 58, pp. 1655–1662, June 2011.

[60] Y. Shi and G. Niu, "Vertical profile design and transit time analysis of nano-scale SiGe HBTs for Terahertz fT," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 213–216, Sept 2004.

[61] B. G. Malm, T. Johansson, T. Arnborg, H. Norstrm, J. V. Grahn, and M. stling, "Implanted collector profile optimization in a SiGe {HBT} process," *Solid-State Electronics*, vol. 45, no. 3, pp. 399 – 404, 2000.

[62] J.-S. Rieh, D. Greenberg, A. Stricker, and G. Freeman, "Scaling of SiGe Heterojunction Bipolar Transistors," *Proceedings of the IEEE*, vol. 93, pp. 1522–1538, Sept 2005.

[63] A. Fischer, H.-J. Osten, and H. Richter, "An equilibrium model for buried SiGe strained layers," *Solid-State Electronics*, vol. 44, no. 5, pp. 869 – 873, 2000.

[64] J. J. H. van den Biesen, "A simple regional analysis of transit times in bipolar transistors," *Solid-State Electronics*, vol. 29, no. 5, pp. 529 – 534, 1986.

[65] M. Bellini, J. Cressler, M. Turowski, G. Avenier, A. Chantre, and P. Chevalier, "3-D Regional Transit Time Analysis of SiGe HBTs on Thin-Film SOI," in *Proceedings of the ECS-SiGe Meeting*, pp. 1079 – 1088, 2008.

[66] K. A. Moen, J. Yuan, P. S. Chakraborty, M. Bellini, J. D. Cressler, H. Ho, H. Yasuda, and R. Wise, "Improved 2-D regional transit time analysis for optimized scaling of SiGe HBTs," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 257–260, Oct 2010.

[67] S. J. Horst, P. Chakraborty, P. Saha, J. D. Cressler, H. Gustat, B. Heinemann, G. G. Fischer, D. Knoll, and B. Tillack, "A comparison of npn vs. pnp SiGe HBT oscillator phase noise performance in a complementary SiGe platform," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 13–16, Oct 2010.

[68] Y. Lu, R. Krithivasan, W.-M. L. Kuo, X. Li, J. D. Cressler, H. Gustat, and B. Heinemann, "A 70 MHz - 4.1 GHz 5th-Order Elliptic g_{m}-C Low-Pass Filter in Complementary SiGe Technology," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 1–4, Oct 2006.

[69] J.-S. Rieh, B. Jagannathan, D. Greenberg, G. Freeman, and S. Subbanna, "A doping concentration-dependent upper limit of the breakdown voltagecutoff frequency product in Si bipolar transistors," *Solid-State Electronics*, vol. 48, no. 2, pp. 339 – 343, 2004.

[70] M. Racanelli and P. Kempf, "SiGe BiCMOS technology for communication products," in *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 331–334, Sept 2003.

[71] L. K. Nanver, N. Nenadovic, V. d'Alessandro, H. Schellevis, H. W. van Zeijl, R. Dekker, D. B. de Mooij, V. Zieren, and J. W. Slotboom, "A back-wafer contacted silicon-on-glass integrated bipolar process. Part I. The conflict electrical versus thermal isolation," *IEEE Transactions on Electron Devices*, vol. 51, pp. 42–50, Jan 2004.

[72] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77 – 89, 1977.

[73] S. M. Sze, *Physics of Semiconductor Devices*. Boston: John Wiley & Sons, 1981.

[74] R. A. Wachnik, T. J. Bucelot, and G. P. Li, "Degradation of bipolar transistors under high current stress at 300 K," *Journal of Applied Physics*, vol. 63, pp. 4734–4740, May 1988.

[75] M. Asheghi, K. Kurabayashi, R. Kasnavi, and K. E. Goodson, "Thermal conduction in doped single-crystal silicon films," *Journal of Applied Physics*, vol. 91, pp. 5079–5088, Apr 2002.

[76] B. Jagannathan, M. Khater, F. Pagette, J.-S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. R. Greenberg, R. Groves, S.-J. Jeng, J. Johnson, E. Mengistu, K. T. Schonenberg, C. M. Schnabel, P. Smith, A. Stricker, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, "Self-aligned SiGe NPN transistors with 285 GHz f_{MAX} and 207 GHz f_T in a manufacturable technology," *IEEE Electron Device Letters*, vol. 23, pp. 258–260, May 2002.

[77] D. C. Ahlgren, M. Gilbert, D. Greenberg, J. Jeng, J. Malinowski, D. Nguyen-Ngoc, K. Schonenberg, K. Stein, R. Groves, K. Walter, G. Hueckel, D. Colavito, G. Freeman, D. Sunderland, D. L. Harame, and B. Meyerson, "Manufacturability demonstration of an integrated SiGe HBT technology for the analog and wireless marketplace," in *Technical Digest of the IEEE International Electron Devices Meeting*, pp. 859–862, Dec 1996.

[78] A. Joseph, D. Coolbaugh, M. Zierak, R. Wuthrich, P. Geiss, Z. He, X. Liu, B. Orner, J. Johnson, G. Freeman, D. Ahlgren, B. Jagannathan, L. Lanzerotti, V. Ramachandran, J. Malinowski, H. Chen, J. Chu, P. Gray, R. Johnson, J. Dunn, S. Subbanna, K. Schonenberg, D. Harame, R. Groves, K. Watson, D. Jadus, M. Meghelli, and A. Rylyakov, "A 0.18 μm BiCMOS technology featuring 120/100 GHz (f_T/f_{MAX}) HBT and ASIC-compatible CMOS using copper interconnect," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 143–146, Sep-Oct 2001.

[79] M. Ieong, P. M. Solomon, S. E. Laux, H.-S. P. Wong, and D. Chidambarrao, "Comparison of raised and Schottky source/drain MOSFETs using a novel tunneling contact model," in *Technical Digest of IEEE International Electron Devices Meeting*, pp. 733–736, Dec 1998.

[80] O. M. Nayfeh, J. L. Hoyt, and D. A. Antoniadis, "Strained-$Si_{1-x}Ge_x/Si$ Band-to-Band Tunneling Transistors: Impact of Tunnel-Junction Germanium Composition and Doping Concentration on Switching Behavior," *IEEE Transactions on Electron Devices*, vol. 56, pp. 2264–2269, Oct 2009.

[81] J. M. C. Stork and R. D. Isaac, "Tunneling in base-emitter junctions," *IEEE Transactions on Electron Devices*, vol. 30, pp. 1527–1534, Nov 1983.

[82] U. Gogineni, J. D. Cressler, G. Niu, and D. L. Harame, "Hot electron and hot hole degradation of UHV/CVD SiGe HBT's," *IEEE Transactions on Electron Devices*, vol. 47, pp. 1440–1448, Jul 2000.

[83] J. Zhao, G.-P. Li, K. Y. Liao, M.-R. Chin, and J. Y.-C. Sun, "On the effects of hydrogen in p-n-p transistors under forward current stress in a C-BiCMOS technology," *IEEE Electron Device Letters*, vol. 14, pp. 4–6, Jan 1993.

[84] B. Heinemann, R. Barth, D. Bolze, J. Drews, G. Fischer, A. Fox, O. Fursenko, T. Grabolla, U. Haak, D. Knoll, R. Kurps, M. Lisker, S. Marschmeyer, H. Rucker, D. Schmidt, J. Schmidt, M. Schubert, B. Tillack, C. Wipf, D. Wolansky, and Y. Yamamoto, "SiGe HBT technology with f_T/f_{max} of 300GHz/500GHz and 2.0 ps CML gate delay," in *Technical Digest of IEEE International Electron Devices Meeting*, pp. 30.5.1–30.5.4, Dec 2010.

[85] H. Rucker, B. Heinemann, and A. Fox, "Half-Terahertz SiGe BiCMOS technology," in *Proceedings of the IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, pp. 133–136, Jan 2012.

[86] M. Schroter, G. Wedel, B. Heinemann, C. Jungemann, J. Krause, P. Chevalier, and A. Chantre, "Physical and Electrical Performance Limits of High-Speed SiGeC HBTs - Part I: Vertical Scaling," *IEEE Transactions on Electron Devices*, vol. 58, pp. 3687–3696, Nov 2011.

[87] M. Rodwell, E. Lind, Z. Griffith, A. Crook, S. Bank, U. Singisetti, M. Wistey, G. Burek, and A. C. Gossard, "On the Feasibility of few-THz Bipolar Transistors," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 17–21, Sept 2007.

[88] B. Heinemann, H. Rucker, R. Barth, J. Bauer, D. Bolze, E. Bugiel, J. Drews, K.-E. Ehwald, T. Grabolla, U. Haak, W. Hoppner, D. Knoll, D. Kruger, B. Kuck, R. Kurps, M. Marschmeyer, H. H. Richter, P. Schley, D. Schmidt, R. Scholz, B. Tillack, W. Winkler, D. Wolnsky, H.-E. Wulf, Y. Yamamoto, and P. Zaumseil, "Novel collector design for high-speed SiGe:C HBTs," in *Technical Digest of IEEE International Electron Devices Meeting*, pp. 775–778, Dec 2002.

[89] B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, T. Grabolla, U. Haak, W. Hoppner, D. Kopke, B. Kuck, R. Kurps, S. Marschmeyer, H. H. Richter, H. Rucker, P. Schley, D. Schmidt, W. Winkler, D. Wolansky, H. E. Wulf, and Y. Yamamoto, "A low-parasitic collector construction for high-speed SiGe:C HBTs," in *Technical Digest of IEEE International Electron Devices Meeting*, pp. 251–254, Dec 2004.

[90] F. J. De La Hidalga, M. J. Deen, and E. A. Gutierrez, "Theoretical and experimental characterization of self-heating in silicon integrated devices operating at low temperatures," *IEEE Transactions on Electron Devices*, vol. 47, pp. 1098–1106, May 2000.

[91] J. Yuan, K. A. Moen, J. D. Cressler, H. Rucker, B. Heinemann, and W. Winkler, "SiGe HBT CML Ring Oscillator With 2.3-ps Gate Delay at Cryogenic Temperatures," *IEEE Transactions on Electron Devices*, vol. 57, pp. 1183–1187, May 2010.

[92] B. A. Orner, Q. Z. Liu, B. Rainey, A. Stricker, P. Geiss, P. Gray, M. Zierak, M. Gordon, D. Collins, V. Ramachandran, W. Hodge, C. Willets, A. Joseph, J. Dunn, J.-S. Rieh, S.-J. Jeng, E. Eld, G. Freeman, and D. Ahlgren, "A 0.13 μm BiCMOS technology featuring a 200/280 GHz (f_T/f_{max}) SiGe HBT," in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 203–206, Sept 2003.

VITA

Partha Sarathi Chakraborty was born in India. He received the Bachelor of Technology degree with a major in Electrical Engineering and a minor in Solid State and Statistical Physics from Indian Institute of Technology (IIT), Kanpur, India; and the Master of Science degree in Electrical Engineering from Arizona State University, Tempe, AZ, USA. He had worked as a Researcher at IIT Delhi; as a PVD Engineer for Applied Materials in Singapore, and as an ESD Engineer at Texas Instruments in Dallas. At Georgia Tech, he has worked with the SiGe Research team on Design, Modeling, Scaling, Characterization, and Reliability of high-performance mixed-signal devices like SiGe HBTs, Graphene FETs, CMOS FETs, etc. While at Georgia Tech, he interned with IBM Microelectronics at Essex Junction, VT. In his own time, he has regularly volunteered locally with Medshare International, Habitat for Humanity, Hands on Atlanta, Discovery Program, First Lego League, etc. He has served as an executive for on-campus honor societies Gamma Beta Phi and Eta Kappa Nu. Partha was the recipient of a scholarship award to attend the 2009 NCN Summer School at Purdue University; the Fall 2014 Georgia Tech Gamma Beta Phi Scholarship Award for Involvement, and the 2015 Roger P. Webb Outstanding Service to Georgia's Community Award by the School of Electrical and Computer Engineering at Georgia Tech. Partha is a founding member of Gradworks, the graduate student wing of the India Club at Georgia Tech. Partha had also conducted trivia quiz events at Georgia Tech. He had regularly volunteered to run bi-weekly grocery shuttles for international students living in the home park area. At Georgia Tech, Partha had served in the library graduate advisory board, as a reviewer for the undergraduate research journal "The Tower", as a judge for the President's Undergraduate Research Award and the Undergraduate Research Symposium, as a mentor and judge for the SURE program in engineering. He had also served as a reviewer for peer-reviewed international journals.

AUTHOR'S PUBLICATIONS

Journals

1. P.S. Chakraborty, A.S. Cardoso, B.R. Wier, A.P. Omprakash, J.D. Cressler, M. Kaynak, and B. Tillack, "A 0.8 THz f_{MAX} SiGe HBT Operating at 4.3 K," *IEEE Electron Device Letters*, 2014, vol. 35, no. 2, pp 151-153.

2. P.S. Chakraborty, K.A. Moen, J.D. Cressler, "An Investigation on the Optimization and Scaling of Complementary SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 60, no. 1, pp 34-41, Jan 2013.

3. A.C. Appaswamy, P.S. Chakraborty, J.D. Cressler, "Influence of Interface Traps on the Temperature Sensitivity of MOSFET mismatch," *IEEE Electron Device Letters*, May 2010.

4. K.A. Moen, P.S. Chakraborty, U. Raghunathan, J.D. Cressler, H. Yasuda, "Predictive Physics-Based TCAD Modeling of the Mixed-Mode Degradation Mechanism in SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 59, no. 11, pp 2895-2901, Nov 2012.

5. A. S. Cardoso, P. S. Chakraborty, N. Karaulac, D. M. Fleischhauer, N. E. Lourenco, Z. E. Fleetwood, A. P. Omprakash, T. D. England, S. Jung, L. Najafizadeh, N. J.-H. Roche, J. H. Warner, A. Khachatrian, D. McMorrow, S. P. Buchner, E. X. Zhang, C. X. Zhang, M. McCurdy, R. A. Reed, D. M. Fleetwood, P. Amouzou, and J. D. Cressler, "Single Event Transient and Total Dose Response of Precision Voltage Reference Circuits Designed in 90 nm SiGe BiCMOS Technology," *IEEE Transactions on Nuclear Science*, Dec. 2014.

6. A. S. Cardoso, P. S. Chakraborty, N. E. Lourenco, T. D. England, P. Saha, D. C. Howard, D. M. Fleischhauer, J. H. Warner, D. McMorrow, S. P. Buchner, P. Paki-Amouzou, T. K. Thrivikraman, and J. D. Cressler, "Evaluating the Effects of Single Event Transients in FET-based Single-Pole Double-Throw RF Switches," *IEEE Transactions on Nuclear Sciences*, vol. 61, no. 2, pp. 756-765, April 2014

7. A. S. Cardoso, P.S. Chakraborty, N.E. Lourenco, P. Song, T.D. England, E.W. Kenyon, N. Karaulac, J. D. Cressler, "Total Ionizing Dose Response of Triple-Well FET-Based Wideband, High-Isolation RF Switches in a 130 nm SiGe BiCMOS Technology," *IEEE Transactions on Nuclear Sciences*, vol.60, no.4, pp. 2567-2573, Aug. 2013

8. Zelei Guo, Rui Dong, P. S. Chakraborty, Nelson Lourenco, James Palmer, Yike Hu, Ming Ruan, John Hankinson, Jan Kunc, John D. Cressler, Claire Berger, and Walt A. de Heer, "High Maximum Oscillation Frequency in Epitaxial Graphene FETs," *Nano Letters*, vol. 13, no. 3, pp 942-947, March 2013.

9. A. S. Cardoso, A. P. Omprakash, P. S. Chakraborty, N. Karaulac, D. M. Fleischhauer, A. Ildefonso, T. G. Bantu, M. A. Oakley, N. E. Lourenco, and J. D. Cressler "On the Cryogenic RF Linearity of SiGe HBTs in a 4th-Generation, 90 nm SiGe BiCMOS Technology," accepted for *IEEE Transactions on Electron Devices*, Jan. 2015

10. J. Hicks, R. Arora, E. Kenyon, P.S. Chakraborty, H. Tinkey, J. Hankinson, C. Berger, W.A. deHeer, E.H. Conrad, J.D. Cressler, "X-ray radiation effects in multilayer epitaxial graphene," *Applied Physics Letters*, vol. 99, no. 23, 2011.

11. N. E. Lourenco, Z. E. Fleetwood, S. Jung, A. S. Cardoso, P. S. Chakraborty, T. D. England, N. J.-H. Roche, A. Khachatrian, D. McMorrow, S. P. Buchner, J. S. Melinger, J. H. Warner, P. Paki, M. Kaynak, B. Tillack, D. Knoll, and J. D. Cressler, "On the Transient Response of Best-of-Breed Complementary (npn + pnp) SiGe HBT BiCMOS Technology," *IEEE Transactions on Nuclear Science*, Dec. 2014

Conferences

1. P.S. Chakraborty, A.C. Appaswamy, P.K. Saha, N.K. Jha, J.D. Cressler, H. Yasuda, B. Eklund, R. Wise, "Mixed-Mode Stress Degradation Mechanisms in pnp SiGe HBTs," *IEEE International Reliability Physics Symposium*, Apr. 2009.

2. P.S. Chakraborty, K.A. Moen, J.D. Cressler, "Predictive TCAD Modeling of the Scaling-Induced, Reverse-Biased, Emitter-Base Tunneling Current in SiGe HBTs," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, Oct. 2011.

3. P.S. Chakraborty, S.J. Horst, K.A. Moen, M. Bellini, J.D. Cressler, "An Investigation of Electro-thermal Instabilities in 150 GHz SiGe HBTs Fabricated on SOI," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, Oct. 2010.

4. P.S. Chakraborty, K.A. Moen, J.D. Cressler, H. Ho, H. Yasuda, B. Eklund, R. Wise, "Predictive Output Conductance Modeling of SiGe HBTs Using TCAD," *International SiGe Technology and Device Meeting*, May 2010.

5. P.S. Chakraborty, K. Moen, M. Bellini, J.D. Cressler, "Investigation of the Device Design Challenges and Optimization Issues Associated with Complementary SiGe HBT Scaling," *IEEE International Semiconductor Device Research Symposium*, Dec. 2009.

6. A. S. Cardoso, P. S. Chakraborty, A. P. Omprakash, N. Karaulac, P. Saha, and J. D. Cressler, "On the Cryogenic Performance of Ultra-Low-Loss, Wideband SPDT RF Switches Designed in a 180 nm SOI-CMOS Technology," *IEEE SOI-3D-Subthreshold Microelectronics Technology*, July 2014.

7. A. S. Cardoso, P. S. Chakraborty, N. E. Lourenco, David M. Fleischhauer, Z. E. Fleetwood, T. D. England, S. Jung, L. Najafizadeh, J. H. Warner, D. McMorrow, S. P. Buchner, E. X. Zhang, C. X. Zhang, M.l McCurdy, R. A. Reed, D. M. Fleetwood,

P. Amouzou, and J. D. Cressler, "Single Event Transient and Total Dose Response of Precision Voltage Reference Circuits Designed in 90 nm SiGe BiCMOS Technology," *IEEE Nuclear and Space Radiation Effects Conference*, July 2014.

8. A. S. Cardoso, P. S. Chakraborty, N. E. Lourenco, T. D. England, P. Saha, D. C. Howard, D. M. Fleischhauer, J. H. Warner, D. McMorrow, S. P. Buchner, P. Paki-Amouzou, T. K. Thrivikraman, and J. D. Cressler, "Evaluating the Effects of Single Event Transients in FET-based Single-Pole Double-Throw RF Switches," *IEEE Nuclear and Space Radiation Effects Conference*, July 2013.

9. A. S. Cardoso, P.S. Chakraborty, N.E. Lourenco, P. Song, T.D. England, E.W. Kenyon, N. Karaulac, J. D. Cressler, "Total Ionizing Dose Response of FET-based Wideband, High-Isolation RF Switches in SiGe BiCMOS Technology," *IEEE Radiation Effects on Components and Systems (RADECS) Conference*, Sept. 2012

10. N. E. Lourenco, Z. E. Fleetwood, P. S. Chakraborty, T. D. England, A.S. Cardoso, S. Jung, N. J.-H. Roche, A. Khachatrian, D. McMorrow, S. P. Buchner, J. S. Melinger, J. H. Warner, P. Paki-Amouzou, M. Kaynak, B. Tillack, D. Knoll, J. D. Cressler, "On the Transient Response of Best-of-Breed Complementary (npn + pnp) SiGe HBT BiCMOS Technology," *IEEE Nuclear and Space Radiation Effects Conference*, July 2014

11. A. S. Cardoso, P. Saha, P. S. Chakraborty, D. M. Fleischhauer, and J. D. Cressler, "Low-loss, Wideband SPDT Switches and Switched-Line Phase Shifter in 180-nm RF CMOS on SOI Technology," *IEEE Radio Wireless Symposium*, pp. 22-24, Jan. 2014

12. Zelei Guo, Rui Dong, P.S. Chakraborty, Nelson Lourenco, James Palmer, Yike Hu, Ming Ruan, John Hankinson, Jan Kunc, John D. Cressler, Claire Berger, and Walt A. de Heer, "Epitaxial Graphene on SiC for Ultra-high Frequency Transistors," *APS*

March Meeting, March 2013.

13. M. Ruan, Y. Hu, Z. Guo, R. Dong, J. Palmer, J. Hankinson, C. Berger, W.A. de Heer, P.S. Chakraborty, N. Lourenco, and J. Cressler, "Structured epitaxial graphene on SiC," *IEEE International Semiconductor Conference Dresden-Grenoble*, September 2012.

14. U.S. Raghunathan, P.S. Chakraborty, B. Wier, J.D. Cressler, H. Yasuda, and P. Menz, "TCAD Modeling of Accumulated Damage During Time-Dependent Mixed-Mode Stress," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, October 2013.

15. S.J. Horst, P.S. Chakraborty, P.K. Saha, J.D. Cressler, H. Gustat, B. Heinemann, G.G. Fischer, D. Knoll, B. Tillack, "A Comparison of *npn* vs. *pnp* SiGe HBT Oscillator Phase Noise Performance in a Complementary SiGe Platform," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting, October 2010*.

16. K.A. Moen, J.H. Yuan, P.S. Chakraborty, M. Bellini, J.D. Cressler, H. Ho, H. Yasuda, R. Wise, "2-D Regional Transit Time Analysis of Scaled SiGe HBTs," *Compact Modeling for RF Workshop, IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, October 2010.

17. J.H. Yuan, J.D. Cressler, K.A. Moen, P.S. Chakraborty, "An Investigation of Collector-Base Transport in SiGe HBTs Designed for Half-Terahertz Speeds," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, October 2010.

18. P. Cheng, C. Grens, A. Appaswamy, P. Chakraborty, J. Cressler, "Modeling Mixed-Mode DC and RF Stress in SiGe HBT Power Amplifiers," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, October, 2008.

19. A.C. Appaswamy, P.S. Chakraborty, J.D. Cressler, "Cryogenic Matching Performance of 90 nm MOSFETs," *IEEE International Semiconductor Device Research Symposium*, December, 2009.

20. B.R. Wier, U.S. Raghunathan, P.S. Chakraborty, J.D. Cressler, H. Yasuda, and P. Menz, "Base Current Degradation Mechanisms in NPN SiGe HBTs subjected to High Current Stress," *IEEE International Semiconductor Device Research Symposium*, December, 2013.

21. M.A. Oakley, B.R. Wier, U.S. Raghunathan, P.S. Chakraborty, and J.D. Cressler, "On the Reliability of SiGe HBT Cascode Driver Amplifiers," *IEEE Radio Frequency Integrated Ciruits Conference*, June, 2014.

22. Z. Xu, G. Niu, L. Luo, P.S. Chakraborty, P. Cheng, D. Thomas, J.D. Cressler, "Cryogenic RF Small-Signal Modeling and Parameter Extraction of SiGe HBTs," *IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, January, 2009.

23. L. Luo, Z. Xu, G. Niu, P.S. Chakraborty, P. Cheng, D. Thomas, K. Moen, J.D. Cressler, M. Mudholkar, H.A. Mantooth, "Wide Temperature Range Compact Modeling of SiGe HBTs for Space Applications," *IEEE Southeastern Symposium on System Theory*, March 2011.

24. U. Raghunathan, P.S. Chakraborty, B. Wier, and J.D. Cressler, "Identifying SOA Boundaries in SiGe HBTs Through Mixed-Mode Device Degradation," *Semiconductor Research Corporation TECHCON*, September 2013.

Book Chapter

1. Partha S Chakraborty, and John D. Cressler, "Hot-Carrier Degradation in Silicon-Germanium Heterojunction Bipolar Transistors," in Tibor Grasser ed., *Hot Carrier Degradation in Semiconductor Devices*, Springer International Publishing, Switzerland, 2015.